

Middlesex University Research Repository

An open access repository of

Middlesex University research

<http://eprints.mdx.ac.uk>

Orengo, Viviane Moreira (2004) Assessing relevance using automatically translated documents for cross-language information retrieval. PhD thesis, Middlesex University. [Thesis]

Final accepted version (with author's formatting)

This version is available at: <https://eprints.mdx.ac.uk/13606/>

Copyright:

Middlesex University Research Repository makes the University's research available electronically.

Copyright and moral rights to this work are retained by the author and/or other copyright owners unless otherwise stated. The work is supplied on the understanding that any use for commercial gain is strictly forbidden. A copy may be downloaded for personal, non-commercial, research or study without prior permission and without charge.

Works, including theses and research projects, may not be reproduced in any format or medium, or extensive quotations taken from them, or their content changed in any way, without first obtaining permission in writing from the copyright holder(s). They may not be sold or exploited commercially in any format or medium without the prior written permission of the copyright holder(s).

Full bibliographic details must be given when referring to, or quoting from full items including the author's name, the title of the work, publication details where relevant (place, publisher, date), pagination, and for theses or dissertations the awarding institution, the degree type awarded, and the date of the award.

If you believe that any material held in the repository infringes copyright law, please contact the Repository Team at Middlesex University via the following email address:

eprints@mdx.ac.uk

The item will be removed from the repository while any claim is being investigated.

See also repository copyright: re-use policy: <http://eprints.mdx.ac.uk/policies.html#copy>

Middlesex University Research Repository:

an open access repository of
Middlesex University research

<http://eprints.mdx.ac.uk>

Orengo, Viviane Moreira, 2004.
Assessing relevance using automatically translated documents for cross-
language information retrieval.
Available from Middlesex University's Research Repository.

Copyright:

Middlesex University Research Repository makes the University's research available electronically.

Copyright and moral rights to this thesis/research project are retained by the author and/or other copyright owners. The work is supplied on the understanding that any use for commercial gain is strictly forbidden. A copy may be downloaded for personal, non-commercial, research or study without prior permission and without charge. Any use of the thesis/research project for private study or research must be properly acknowledged with reference to the work's full bibliographic details.

This thesis/research project may not be reproduced in any format or medium, or extensive quotations taken from it, or its content changed in any way, without first obtaining permission in writing from the copyright holder(s).

If you believe that any material held in the repository infringes copyright law, please contact the Repository Team at Middlesex University via the following email address:
eprints@mdx.ac.uk

The item will be removed from the repository while any claim is being investigated.

Assessing Relevance Using Automatically Translated Documents for Cross-Language Information Retrieval

**A thesis submitted to Middlesex University in partial fulfilment of the
requirements for the degree of Ph.D.**

March 2004.



Viviane Moreira Orengo

Acknowledgements

I would like to thank the following people for their help with this thesis.

- Dr. Christian Huyck, for the excellent supervision and constant motivation provided throughout this work.
- Dr. Michael Wing, for the advice and for enabling me to start off this PhD.
- Dr. Mounia Lalmas, for the helpful comments.
- Ana Luísa Barros for the invaluable help with the user experiment.
- My parents, Paulo and Tânia, and my brother Henrique for the support and encouragement.
- My husband, Giuliano, for his love and support.
- My friends, Joana, Silvana, Victor, Rosane and Andrea for their friendship and company.
- The staff from Computing Science Research Office for their constant support.

ABSTRACT

This thesis focuses on the Relevance Feedback (RF) process, and the scenario considered is that of a Portuguese-English Cross-Language Information Retrieval (CLIR) system. CLIR deals with the retrieval of documents in one natural language in response to a query expressed in another language. RF is an automatic process for query reformulation. The idea behind it is that users are unlikely to produce perfect queries, especially if given just one attempt. The process aims at improving the query specification, which will lead to more relevant documents being retrieved. The method consists of asking the user to analyse an initial sample of documents retrieved in response to a query and judge them for relevance.

In that context, two main questions were posed. The first one relates to the user's ability in assessing the relevance of texts in a foreign language, texts hand translated into their language and texts automatically translated into their language. The second question concerns the relationship between the accuracy of the participant's judgements and the improvement achieved through the RF process.

In order to answer those questions, this work performed an experiment in which Portuguese speakers were asked to judge the relevance of English documents, documents hand-translated to Portuguese, and documents automatically translated to Portuguese. The results show that machine translation is as effective as hand translation in aiding users to assess relevance. In addition, the impact of misjudged documents on the performance of RF is overall just moderate, and varies greatly for different query topics.

This work advances the existing research on RF by considering a CLIR scenario and carrying out user experiments, which analyse aspects of RF and CLIR that remained unexplored until now. The contributions of this work also include: the investigation of CLIR using a new language pair; the design and implementation of a stemming algorithm for Portuguese; and the carrying out of several experiments using Latent Semantic Indexing which contribute data points to the CLIR theory.

TABLE OF CONTENTS

1	INTRODUCTION	1
1.1	Scope and Objectives of this Research.....	1
1.2	Thesis Structure	3
2	LITERATURE REVIEW	4
2.1	Information Retrieval.....	4
2.1.1	Classic Models	6
2.1.1.1	Boolean Model	7
2.1.1.2	Vector Space Model.....	7
2.1.1.3	Probabilistic Model	8
2.1.2	Evaluation in Information Retrieval	10
2.2	Cross-Language Information Retrieval.....	14
2.2.1	Related Terms	15
2.2.2	The Problem of CLIR	16
2.2.3	Brief History	17
2.2.4	Areas of Application.....	17
2.2.5	The User of CLIR Systems.....	18
2.2.6	Evaluation of CLIR Systems	19
2.2.7	Approaches to CLIR.....	20
2.2.7.1	Machine Translation	20
2.2.7.2	Thesauri	21
2.2.7.3	Dictionaries	22
2.2.7.4	Corpus-based Techniques.....	23
2.2.7.5	Current State of the Art	24
2.3	Latent Semantic Indexing.....	25
2.3.1	Method.....	25
2.3.2	Example	27
2.3.3	Other applications.....	32
2.3.4	Discussion of LSI	33
2.4	CLIR Using Latent Semantic Indexing.....	34
2.5	Relevance Feedback	36
2.5.1	Brief History	37
2.5.2	Relevance Feedback and the Different IR Models.....	38
2.5.2.1	Relevance Feedback and the Vector-Space Model	38
2.5.2.2	Relevance Feedback and the Boolean Model.....	39
2.5.2.3	Relevance Feedback and the Probabilistic Model	41
2.5.2.4	Relevance Feedback and Latent Semantic Indexing	42
2.5.2.5	Other Methods	43
2.5.3	Evaluation	44
2.5.4	Discussion of Relevance Feedback.....	45
2.6	Summary of the Chapter	47
3	A STEMMING ALGORITHM FOR PORTUGUESE.....	48
3.1	Introduction	48
3.2	The Algorithm	49
3.3	Difficulties in Stemming Portuguese.....	53
3.4	Evaluating The Stemmer	55
3.4.1	Vocabulary Reduction	55
3.4.2	Comparison Against Expected Output	56
3.4.3	Paice's Evaluation Method.....	56

3.4.3.1	Method Description	57
3.4.3.2	Comparison between the Two Stemmers.....	60
3.5	Summary of the Chapter	60
4	PORTUGUESE-ENGLISH CROSS-LANGUAGE INFORMATION RETRIEVAL.....	62
4.1	Why Portuguese?	62
4.2	The CLIR System	63
4.3	Corpora	64
4.4	Experiments	64
4.4.1	Weighting Scheme.....	66
4.4.2	Number of dimensions.....	67
4.4.3	Monolingual vs Bilingual Retrieval	68
4.4.4	Effects of Stemming	70
4.4.5	Similarity between words	72
4.5	Discussion.....	73
4.6	Summary of Chapter.....	73
5	RELEVANCE FEEDBACK AND CROSS-LANGUAGE INFORMATION RETRIEVAL.....	75
5.1	Background	76
5.2	Choosing an RF Method.....	77
5.3	Experimental Design.....	80
5.3.1	Searcher	80
5.3.2	Document Collection and System.....	81
5.3.3	Query Topics.....	81
5.3.4	Procedure	83
5.4	Users Ability in Making Relevance Judgements	86
5.4.1	Number of Mistakes.....	86
5.4.2	Overlap	87
5.4.3	Difficulty, Confidence and Knowledge.....	88
5.4.4	Testing Other User Groups.....	91
5.4.4.1	Testing Bilingual Users.....	91
5.4.4.2	Testing Native Speakers	92
5.4.5	Comparison Among Groups	92
5.5	Errors in Judgement and Change in Performance.....	93
5.5.1	Evaluation by System	97
5.5.2	Evaluation by Topic.....	98
5.6	Findings	99
5.6.1	RF improves the performance of a CLIR-LSI system	99
5.6.2	Machine Translation is as Effective as Hand Translation.....	100
5.6.3	Misjudged Documents have a Moderate Effect	102
5.6.4	Difficulty of Task and Accuracy in Judgement.....	103
5.6.5	Proper Names help the RF process	104
5.6.6	Participants Feedback.....	105
5.7	Summary of Chapter.....	105
6	DISCUSSION	107
6.1	What is MT Good For?	107
6.2	Addressing Common Criticisms of the Methodology.....	108
6.2.1	Validity of Relevance Assessments.....	108
6.2.2	CLEF Topics and Relevance Assessments	112
6.2.3	The Use of Ready Query Topics.....	112
6.3	RF in the Real World.....	113

6.4	Algorithmic Considerations	114
6.4.1	LSI and Large Scale IR	114
6.4.2	Using Negative Feedback.....	115
6.5	Generalising from the Results	115
7	CONTRIBUTIONS AND FUTURE WORK.....	117
7.1	Contributions of the Thesis.....	117
7.1.1	CLIR was extended to a new language pair	117
7.1.2	Development of a Portuguese stemming algorithm	118
7.1.3	Running of several LSI - CLIR experiments	118
7.1.4	Analysis of translated texts for relevance assessment.....	118
7.1.5	Study of the relationship between accuracy in judgement and performance improvement.....	119
7.2	Future Work.....	119
7.2.1	Variations on RF Experiment.....	119
7.2.2	Analyse the characteristics of the topics	119
7.2.3	Propose a method for Negative Feedback	120
7.2.4	Improving the Portuguese stemmer	120
7.2.5	Use comparable corpora for training the IR system.....	120
7.3	Final Word.....	121
	REFERENCES.....	122
	APPENDIX A – PORTUGUESE STEMMING RULES	136
	APPENDIX B – RELEVANCE FEEDBACK QUESTIONNAIRE	143
	APPENDIX C – RELEVANCE FEEDBACK QUESTIONNAIRE (ENGLISH SUBJECTS).....	208
	APPENDIX D – CLEF OFFICIAL JUDGEMENTS	229
	APPENDIX E – DATA FROM USER EXPERIMENTS	230

LIST OF FIGURES

Figure 2.1 – Information Retrieval Process.....	5
Figure 2.2 – Example of Boolean Retrieval	7
Figure 2.3 – Example of Retrieval in the Vector-Space Model.....	8
Figure 2.4 – IR Example.....	10
Figure 2.5 – Precision at 11 recall points	12
Figure 2.6– Singular Value Decomposition	26
Figure 2.7 – SVD Example.....	29
Figure 2.8 – Dimension Reduction.....	30
Figure 2.9 – Plot of terms and documents using two dimensions.....	30
Figure 2.10 – Bilingual term by document matrix	35
Figure 2.11 Relevance Feedback Process	37
Figure 3.1 – Sequence of steps for the Portuguese Stemmer	50
Figure 4.1– Languages in number of native speakers [Ethnologue].....	63
Figure 4.2 – Web hosts by language [<i>Internet Software Consortium</i>]	63
Figure 4.3 – Sample dual-language document before stemming	65
Figure 4.4 – Sample dual-language document after stemming	65
Figure 4.5 – Number of dimensions	67
Figure 4.6 – Comparison between Monolingual and Bilingual IR.....	69
Figure 4.7 – Recall-Precision Curves for Stemming Experiment.....	71
Figure 5.1 – Performance of RF methods for LSI	79
Figure 5.2 – English knowledge question.....	81
Figure 5.3 – Questions on difficulty, confidence and familiarity	89
Figure 5.4 – Histogram for Difficulty of the Task (per system)	89
Figure 5.5 – Histogram for Confidence in Judgements (per system).....	89
Figure 5.6 – Histograms for Prior Knowledge of the Topics.....	90
Figure 5.7 – Histogram for change in precision (by user)	95
Figure 5.8 – Recall–Precision curves using the residual collection.	95

LIST OF TABLES

Table 2.1 – Sample collection of documents.....	27
Table 2.2 – Term-by-document matrix A	28
Table 2.3 – Term-term comparisons	31
Table 2.4 – Document-document comparisons.....	31
Table 2.5 – Query-document comparison.....	32
Table 2.6 – Matrix A ₂ – Approximation of matrix A using 2 dimensions.....	32
Table 3.1 – Calculating GDMT and GDNT.....	58
Table 3.2 – Calculating GUMT	59
Table 3.3 – Calculating GWMT	59
Table 4.1 – Details of the test collections.....	64
Table 4.2 – Summary of Results for Monolingual-Bilingual Comparison	68
Table 4.3 – Summary of Results for Stemming Experiment.....	70
Table 4.4 – Term similarity across languages.....	72
Table 5.1 – Details of LA Times.....	81
Table 5.2 – Number of relevant documents per topic.....	84
Table 5.3 – Subject matrix showing topic-system combination and the presentation order	85
Table 5.4 – Summary of Judgement Mistakes	87
Table 5.5 – Overlap Example.....	88
Table 5.6 – Overlap by System	88
Table 5.7 – Average precision figures for initial and feedback runs.....	94
Table 5.8 – Results for Bilingual Users	91
Table 5.9 – Results Native Speakers	92
Table 5.10 – Comparison among groups	93
Table 5.11 – Change in average precision (by topic)	100
Table 5.12 – Translation Errors.....	101
Table 5.13 – Participants ranked by change and overlap	96
Table 5.14 – Correlation between change in precision and accuracy in judgement	96
Table 5.15 – Correlation between change in precision and accuracy in judgement (by system) .	97
Table 5.16 – Correlation between change in precision and accuracy in judgement (by topic).....	98
Table 6.1 – Some factors that may affect the judgement of relevance.....	110
Table 6.2 – Ranking of Text-Handling Tasks [Taylor 1998]	108

LIST OF ABBREVIATIONS

This list contains all abbreviations used throughout the thesis, along with their meanings, and the page in which they have been defined.

CLEF	Cross-Language Evaluation Forum, page 13
CLIR	Cross-Language Information Retrieval, page 1
ECDL	European Conference on Digital Libraries, page 19
IR	Information Retrieval, page 1
LSI	Machine Translation, page 2
MLIA	Multilingual Information Access, page 15
MLIR	Multilingual Information Retrieval, page 15
MRD	Machine-Readable Dictionary, page 22
MT	Machine Translation, page 16
NLP	Natural Language Processing, page 17
PRF	Pseudo Relevance Feedback, page 37
RF	Relevance Feedback, page 36
RSLP	"Removedor de Sufixos da Língua Portuguesa", page 48
SVD	Singular Value Decomposition, page 25
SVM	Support Vector Machine, page 43
tf-idf	term frequency - inverse document frequency, page 5
TREC	Text REtrieval Conference, page 13

1 INTRODUCTION

The increasing availability of information in different languages and the growing number of people speaking different mother tongues who want to find information have been motivating research on Cross-Language Information Retrieval (CLIR). CLIR is responsible for receiving search requests expressed in one language and retrieving documents written in another language. The scope of CLIR typically involves mapping a concept from one language into another.

Some traditional Information Retrieval (IR) techniques, such as Relevance Feedback (RF) acquire a new dimension in this cross-linguistic environment. The RF process consists in showing the searcher an initial list of retrieved documents in response to a query. The searcher is asked to judge the documents for relevance. The original query is then modified, based on the user's choices, and resubmitted. The aim is to obtain a better representation of the information need and therefore return a better set of documents. In a CLIR scenario, the user's ability to assess the relevance of documents written in a foreign language, or translated into his native language by some means, becomes an issue.

1.1 Scope and Objectives of this Research

This research began with a general interest in CLIR, specifically on the use of Portuguese queries to retrieve documents written in English. The next step was then to survey methods for implementing a CLIR system, and the method chosen

was Latent Semantic Indexing (LSI). This choice was mainly motivated by the good results achieved by LSI and the straightforward adaptation of the system to do CLIR. In order to carry out experiments, some resources such as stop word lists and stemmers were needed. While it was very easy to find them for English, resources for Portuguese were very scarce. Up to now the only available stemmer for Portuguese was a translated version of a well-known algorithm that had been designed for English. The problem is that there are large differences between the two languages that need to be taken into account. With the aim of addressing this lack, a Portuguese stemming algorithm has been designed and implemented. Several bilingual CLIR experiments were then performed. The result was a reasonable CLIR system. The next step was to look for methods for improving the results. RF was the selected method, and from there several questions began to emerge.

The aim of this thesis is to advance the theory and practice of CLIR with a focus on the RF process. The objectives of this work include:

- *Explore Portuguese-English CLIR.* This work is the first to use Portuguese, which is one of the biggest languages in number of native speakers.
- *Design and make available resources for the processing of Portuguese.* Despite being a major language, Portuguese is very under-represented and linguistic resources are scarce. Therefore, one of the aims of this work is to develop an effective stemming algorithm for Portuguese, which should be made freely available to other researchers.
- *Find out how accurately users can assess the relevance of texts in a foreign language, texts hand translated into their language, and texts automatically translated into their language.* In order to answer those questions, an experiment in which Portuguese speakers are asked to judge documents for relevance was carried out.
- *Find out the relationship between the accuracy of the judgements and the improvement achieved through the RF process.* The objective here is to analyse the

impact that misjudged documents have on the performance improvement that can be achieved with RF.

1.2 Thesis Structure

The remainder of this thesis is organised as follows:

- Chapter 2 presents the literature review, in which basic concepts of IR, CLIR, RF and LSI are exposed.
- Chapter 3 describes the Portuguese stemming algorithm and evaluates its performance.
- Chapter 4 presents experiments done for Portuguese-English CLIR using LSI.
- Chapter 5 contains the RF experiment along with the findings.
- Chapter 6 discusses several issues relating to the experiments performed and the broader context of the field.
- Chapter 7 summarises the contributions and suggests future work.

The thesis also contains four Appendices:

- Appendix A details the Portuguese stemming rules defined by the algorithm proposed in Chapter 3.
- Appendix B presents the RF questionnaire submitted to the Portuguese subjects, along with the topics and documents judged for relevance.
- Appendix C contains the RF questionnaire submitted to the English subjects, along with the topics and documents judged for relevance.
- Appendix D contains the official CLEF judgements for the documents used in the RF experiment.
- Appendix E presents the data collected from the user experiments and the statistical tests performed.

2 LITERATURE REVIEW

This chapter presents some background concepts that will be referred to in later chapters. The first section provides a basic notion of Information Retrieval. The second section on Cross-Language Information Retrieval, which is the theme of this thesis. The third section describes, Latent Semantic Indexing, the technique used in all experiments reported in Chapters 4 and 5. The fourth section explains how Latent Semantic Indexing can be applied to Cross-Language Information Retrieval. Finally, the fifth section is on Relevance Feedback, which is applied in the experiments described in Chapter 5. For a more detailed introduction to each of these subjects, please refer to the relevant texts pointed in each section.

2.1 Information Retrieval

IR deals with the representation, storage, organisation, and access to information items. The representation and organisation of such items should provide the user with easy access to the information in which he is interested [Baeza-Yates 1999].

The user typically needs to translate his information need into a query, which can be processed by the IR system. This query is typically represented by a set of keywords. The output produced by the IR system is a ranked list of the documents that are likely to satisfy the user's information need; this process is depicted in Figure 2.1. IR research is focused on capturing the information provided by the users to improve the results the system returns.

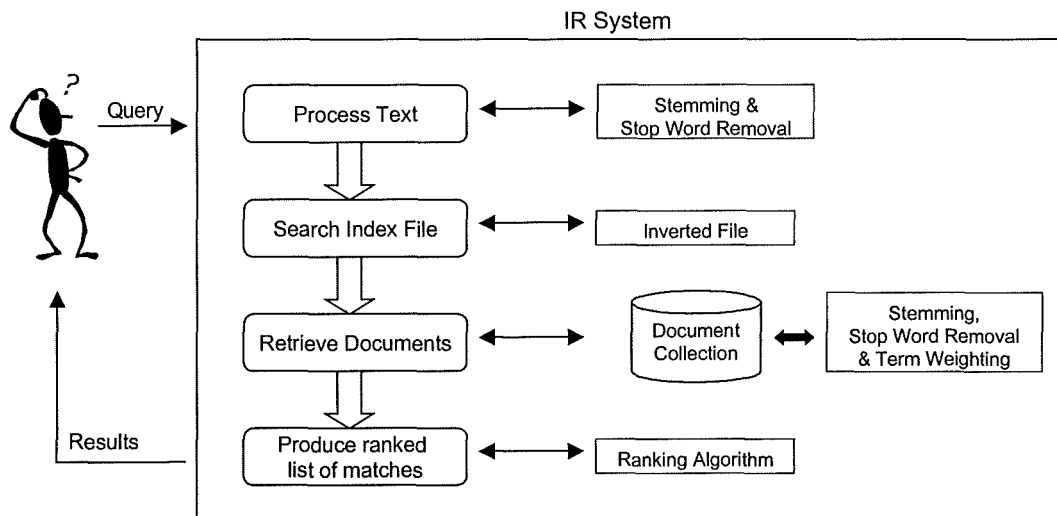


Figure 2.1 – Information Retrieval Process

In most IR systems, the database of documents needs to go through some pre-processing before the documents are made available for retrieval. This typically involves:

- **Stemming:** stemming reduces a word to its root by removing suffixes. For example, the words: "presentation", "presented", "presenting" could all be reduced to a common root "present". This is based on the assumption that posing a query with the term *presenting* implies an interest in documents containing the words *presentation* and *presented*. Stemming considerably reduces the number of index terms and aims at improving recall.
- **Stop word removal:** stop words are the most common words in a language (e.g. conjunctions, articles, prepositions...), they typically consist of about a hundred terms which represent up to 50% of the total words in the documents. Such words carry little meaning and therefore can be ignored, which means they are not included in the index.
- **Term Weighting:** term weights quantify the importance of the terms. The basic idea is to assign low weights to words that appear in many documents and high weights to words that appear in few documents as those are more distinctive. Several term weighting schemes have been proposed. However, the best known method is **tf-idf** (term frequency times inverse document frequency),

which is a function of the number of occurrences of a term in a document and the number of occurrences of that term across the document collection. There are several variations of the tf-idf formula, one formulation is given below:

$$w_{ij} = tf_{ij} \times \log_2 \frac{N}{n} \quad (1)$$

where: w_{ij} = weight of term T_j in document D_i
 tf_{ij} = frequency of term T_j in document D_i
 N = number of documents in the collection
 n = number of documents in which the term T_j appears

This formula benefits large documents, since the term frequency tends to be greater. To avoid this problem, variations of the formula above normalise the document length.

- **Indexing:** an index is a data structure that allows quick searching of the text. The most popular index for text searching is the *inverted file*. An inverted file has two elements: a lexicon (a list of all terms that appear on the database) and a list of documents (or positions) where each term appears. The use of an inverted file dramatically improves the search efficiency. The downside is that the system has the overhead of storing and maintaining this data structure, which can range from 10 percent to over 100 percent of the size of the text itself. For a comprehensive discussion on indexing algorithms, please refer to Frakes and Baeza-Yates [Frakes 1992] or Witten, Moffat and Bell [Witten 1999].

2.1.1 Classic Models

An IR model is concerned with the specification of the representations used for documents and queries, and how they can be compared. Interest in IR started in the 1940s [Frakes 1992], since then several models have been proposed. This section presents an overview of the classic IR models, namely: Boolean, Vector-

Space and Probabilistic. For a more comprehensive explanation please refer to [Baeza-Yates 1999] and [van Rijsbergen 1979].

2.1.1.1 Boolean Model

This is a simple retrieval model based on set theory and Boolean Algebra, which provides an easy to understand framework. It considers that index terms are present or absent in a document. The keywords in the query are combined using the Boolean operators “and”, “or” and “not”. For example, the retrieval for the query: *cat and dog and not bear*, can be represented as shown by Figure 2.2.

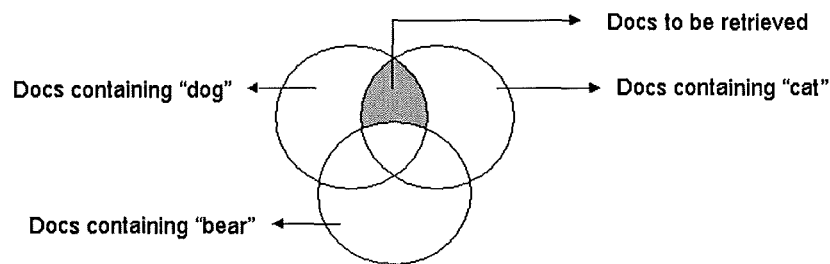


Figure 2.2 – Example of Boolean Retrieval

This model has received great attention in the past, however, it has a major drawback: a document is either classified as relevant or not relevant, so there is no grading or ranking. Despite this shortcoming, the Boolean model is still dominant in commercial systems.

Salton, Fox and Wu [Salton 1983] have proposed an extension to the Boolean model in order to give support to term weighting and ranking. Despite its advantages, the Extended Boolean model has not been used extensively.

2.1.1.2 Vector Space Model

This model, proposed by Salton [Salton 1971b], assigns weights to the terms in the documents and in the queries. The queries and documents are represented by sets of independently weighted terms not connected by Boolean operators. The weights are used to compute the similarity between each document in the

database and the given query. Documents and queries are treated as vectors in a t -dimensional space (where t is equal to the number of terms in the collection) and the degree of similarity between them can be measured, for instance, by the cosine of the angle between the vectors.

Figure 2.3 shows an example of retrieval in the vector-space model. Suppose there are 6 documents in the collection and a query Q is posed. The most similar documents will be retrieved; this is measured by the cosine of the angle between the query and each document. For the example shown in Figure 2.3, the most relevant document is d_2 as the cosine between its vector and the query vector is closer to 1. The most dissimilar (or irrelevant) document is d_5 as the cosine of the angle is very close to -1.

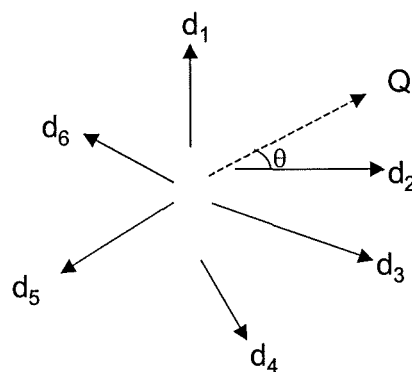


Figure 2.3 – Example of Retrieval in the Vector-Space Model

2.1.1.3 Probabilistic Model

This model was proposed by Robertson and Sparck-Jones [S. Robertson 1976] and is also known as Binary Independence Retrieval. Like Boolean retrieval, the probabilistic model is based on binary document indexing, i.e. the index term is either present or absent from the document. A document is either considered relevant or irrelevant to a query. This model assumes the independence of index terms, which leads to the notion that the relevance of one document reveals nothing about the relevance of another document.

The fundamental idea is: given a user query, there is a set which contains exactly the relevant documents and no other. This set of documents is the ideal answer set. The querying process involves describing the properties of this ideal answer set. Given a query and a document, the probabilistic model tries to estimate the probability that a document is relevant for the query. The documents are then ranked in decreasing order of probability of relevance to the query. The model assumes that the probability of relevance depends on the representation of the query and the document only.

The similarity between a query Q and a document d_j is given by the ratio between the probability that the document d_j is relevant to the query and the probability that d_j is not relevant to the query:

$$\text{sim}(d_j, Q) = \frac{P(\text{rel} | d_j)}{P(\text{nrel} | d_j)} \quad (2)$$

At the beginning of the retrieval process, there is no information on the relevance of the documents, therefore an initial probability guess is made based on the distribution of the words in the documents. A simplifying assumption would be that the probability of relevance is constant (say 0.5) for all index terms. The distribution of index terms among the non-relevant documents can be approximated by their distribution across the whole collection. Thus:

$$\begin{aligned} P(k_i | \text{rel}) &= 0.5 \\ P(k_i | \text{nrel}) &= \frac{n_i}{N} \end{aligned} \quad (3)$$

where $P(k_i | \text{rel})$ is the probability that the index term k_i is present in a relevant document, and $P(k_i | \text{nrel})$ is the probability that the index term is present in a non-relevant document. n_i is the number of documents that contain the index term and N is the total number of documents in the collection.

Once an initial retrieval set is obtained, the guess on the relevance probability can be improved. The probability $P(k_i | rel)$ can be approximated by the distribution of the index term in the documents retrieved so far, and $P(k_i | nrel)$ can be approximated by assuming that all non-retrieved documents are not relevant. This process can be repeated recursively with or without the assistance of a human subject to judge the retrieved documents for relevance.

There is some disagreement as to which method is superior: the vector-space or the probabilistic model. Experiments carried out by different people have arrived at contrasting conclusions. According to Baeza-Yates and Ribeiro-Neto [Baeza-Yates 1999], Croft performed some experiments and suggested that the probabilistic model provides a better retrieval performance. However, experiments done by Salton and Buckley showed that the vector-space model is expected to outperform the probabilistic model with general collections. Both models enjoy popularity among researchers.

2.1.2 Evaluation in Information Retrieval

The most common way to describe retrieval performance is to calculate how many of the relevant documents have been retrieved and how early in the ranking they were listed [Witten 1999]. Figure 2.4 shows an example of an information request for which there is a set of relevant documents and an answer set.

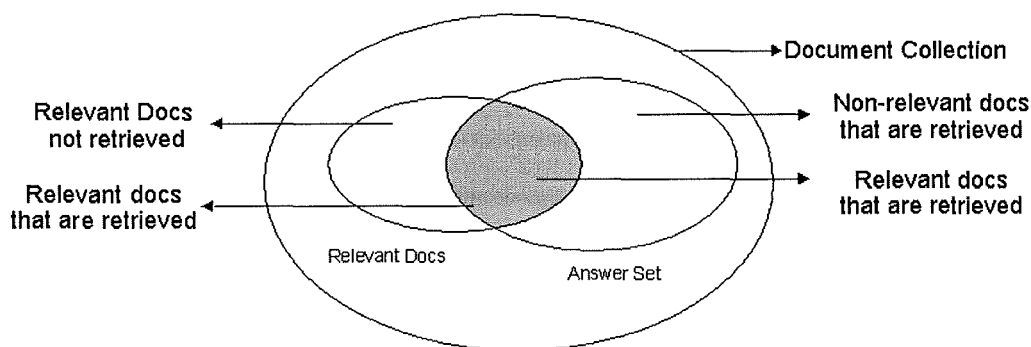


Figure 2.4 – IR Example

The intersection between the two areas represents the documents that were correctly retrieved. As can be seen, some relevant documents were not retrieved and some irrelevant documents were retrieved.

The two traditional measures of performance are recall and precision, which can be defined as follows:

- **Recall** is the fraction of relevant documents that has been retrieved.

$$R = \frac{\text{number relevant that are retrieved}}{\text{total number relevant}} \quad (4)$$

- **Precision** is the fraction of the retrieved documents that is relevant.

$$P = \frac{\text{number relevant that are retrieved}}{\text{total number retrieved}} \quad (5)$$

For example, suppose that 20 documents were retrieved in answer to a given query, 7 of them were relevant, and the total number of relevant documents for that query (correct answer) is 10. Then, the recall would be 70% ($R=7/10$) and the precision would be 35% ($P=7/20$). Typically, a high recall is accompanied by low precision and vice-versa.

The standard IR evaluation experiment takes the values for precision at 11 different recall points (0%, 10%, ..., 100%). Using the data from the example, let us suppose that the ranked list of documents in the answer set is as follows (the relevant documents are marked with an asterisk):

d ₁ *	d ₅	d ₉ *	d ₁₃ *	d ₁₇
d ₂	d ₆ *	d ₁₀	d ₁₄	d ₁₈ *
d ₃	d ₇	d ₁₁ *	d ₁₅	d ₁₉
d ₄ *	d ₈	d ₁₂	d ₁₆	d ₂₀

The first document, which corresponds to 10% of all relevant documents (as there are 10 in total) is relevant, so it can be said that there is 100% precision at 10% recall (1 relevant document out of 1 retrieved). The next relevant document is the

fourth, thus the precision correspondent to 20% recall is 50% (two relevant documents out of 4 retrieved). Proceeding with the analysis, the last relevant document is the eighteenth, at that point the recall is 70% and the correspondent precision is 38% (7 relevant documents out of 18 retrieved). Note that the precision at recall levels over 70% drops to 0 because not all relevant documents were retrieved. For the 0% recall level the precision is determined using an interpolation method that states that the precision at a standard recall level i is the maximum precision obtained for any actual recall level greater than or equal to i . For the data in the example, the precision for the recall level 0% is the maximum known precision, which is 100%. Figure 2.5 shows the chart for precision at 11 recall points for the example.

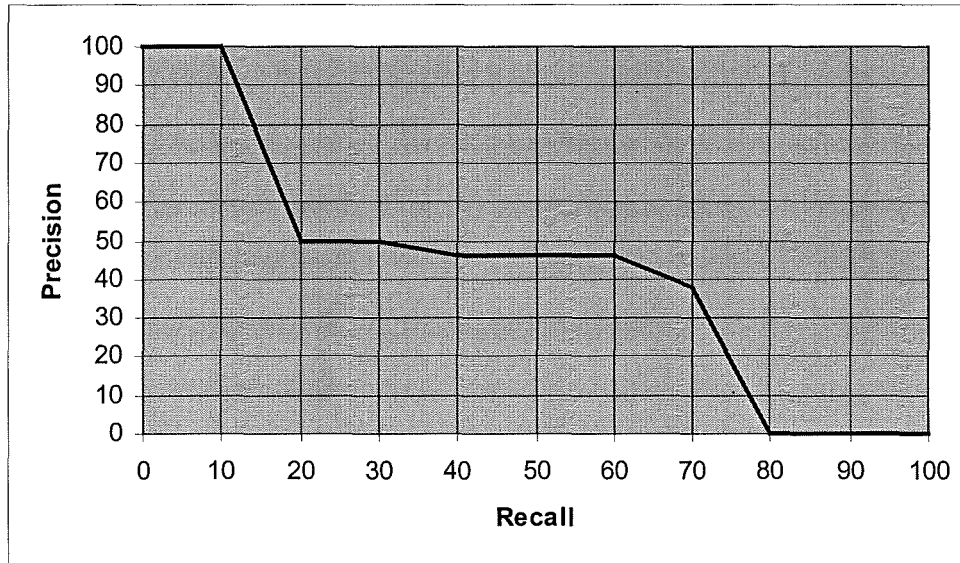


Figure 2.5 – Precision at 11 recall points

The chart in the example was built using the measures for one query only, in order to evaluate the performance of the system over all test queries, the average precision is used. Its formula is given below:

$$\overline{P}(r) = \sum_{i=1}^{N_q} \frac{P_i(r)}{N_q} \quad (6)$$

where $\bar{P}(r)$ is the average precision at the recall point r , N_q is the number of queries used, $P_i(r)$ is the precision at recall level r for the i^{th} query.

The recall-precision chart is used to evaluate ranked lists of retrieved documents and it is also useful for comparing the retrieval performance of different retrieval algorithms. However, in some situations it might be necessary to evaluate an unordered set of retrieved documents, or it might be desirable to obtain a measure that can be expressed as a single value summary. One way of combining precision and recall into a single parameter is the **F-measure** [van Rijsbergen 1979]. It has two important properties: it penalises systems that achieve high recall by sacrificing precision and vice-versa, and because it is a weighted combination, it is possible to choose which measure to give more emphasis. The formula is:

$$F = \frac{(\beta^2 + 1) \times P \times R}{(\beta^2 \times P) + R} \quad (7)$$

If precision and recall should receive the same emphasis, then $\beta = 1$ (or P&R); $\beta=2$ (or P&2R) emphasises recall twice as much as precision and $\beta = 0.5$ (or 2P&R) emphasises precision twice as much as recall. In the example, $P=35\%$ and $R=70\%$, therefore the values for F-measure are $F_{(P\&R)}=46\%$, $F_{(P\&2R)}= 58\%$ and $F_{(2P\&R)}= 38\%$.

Other retrieval measures have been proposed and the problems of recall and precision have been studied by Raghavan, Jung, and Bollman [Raghavan 1989] among others. However, recall and precision are the measures used by major evaluation campaigns such as TREC (Text Retrieval Conference) and CLEF (Cross-Language Evaluation Forum). For that reason, they are used in evaluating the experiments performed for this thesis.

In order to calculate those measures, one needs to know the “correct answers”; this is done using test collections. IR test collections consist of a document

collection, a set of queries and the relevance judgements, which assign for each query which documents are relevant.

Creating an IR test collection is a major task. Even for small collections the assignment of relevance assessments is a big endeavour. The Time collection, for example, has 425 documents and 83 queries, requiring over 35 thousand relevance judgements. For the Cranfield collection each document was assessed by more than one judge; reportedly more than half a million judgements were made. This becomes totally impractical for larger collections. In order to solve this problems, evaluation campaigns e.g. [Kluck 2002; Voorhees 2001] use the **pooling** method. This approach consists in building judgement pools from retrieved documents obtained from the submitted runs from participants of the evaluation campaign. Only the top n documents from selected runs are added to the pool. The relevance assessors will only need to judge the documents from the pool. This technique dramatically reduces the number of assessments that need to be made. The downside is that all unjudged documents are assumed irrelevant, which is not necessarily true. However, evaluations of the TREC collection performed by Zobel [Zobel 1998] show that it is not biased against unjudged runs. The pooling method provides a good basis because the pool will contain documents judged as relevant by several different IR systems. With enough systems and a sufficiently deep pool, there is some likelihood that most relevant documents are going to be found [Zobel 1998].

2.2 Cross-Language Information Retrieval

Cross-Language Information Retrieval (CLIR) is the retrieval of documents in one natural language, based on a query formulated in another natural language, e.g. retrieval of English documents based on a Spanish query.

The main motivation for CLIR is the growing need of exploring documents in foreign languages. This need has increased dramatically with the explosive growth of the Internet, which has made digital resources available to more and

more countries. The physical distance between the user and the information is no longer present, however there is still the language barrier to be surmounted.

2.2.1 Related Terms

In the literature about CLIR there is little agreement on the usage of terms. The expression Multilingual Information Retrieval (MLIR) is frequently used to mean the same as CLIR. However, Hull and Grefenstette [Hull & Grefenstette 1996] have defined MLIR as:

- Information Retrieval in a language other than English. In this case both the query and the document collection are in the same language (monolingual collection, monolingual retrieval);
- Information Retrieval on a parallel document collection or on a multilingual document collection where the search space is restricted by the query language. In this case the documents are in many languages but once the query is entered the language of interest is fixed and only the documents in that language are searched (multilingual documents, monolingual retrieval);
- Information Retrieval on a monolingual collection which can be queried in many languages (monolingual collection, cross-language retrieval);
- Information Retrieval on a multilingual document collection where queries can retrieve documents in many languages (multilingual collection, cross-language retrieval);
- Information Retrieval on multilingual documents, i.e. more than one language can be present in individual documents (multilingual collection, cross-language retrieval).

A new term which has been used by Peters and Sheridan [Peters 2000] is Multilingual Information Access (MLIA). MLIA is used in a broader sense addressing the problem of accessing, querying and retrieving information in any

language at any level of specificity and includes all issues that involve the overall management of multilingual information, such as character encoding, language identification, indexing of collections in multiple languages, etc. There is yet another term used by Frederking [Frederking 1997] and Yang [Yang 1997] as a synonym for CLIR: Translingual Information Retrieval. In this thesis the term CLIR is used because it more clearly concentrates on the process of accepting queries in one language and retrieving relevant documents in other languages.

2.2.2 The Problem of CLIR

According to Grefenstette [Grefenstette 1998], CLIR involves basically three problems:

- The first problem is knowing how a term expressed in one language might be written in another.
- The second problem is deciding which of the possible translations should be retained. Retaining more than one translation is useful in promoting recall.
- The third problem is deciding how to properly weight the importance of translation alternatives when more than one is retained.

The first problem has to do with alternatives for crossing the language barrier. Several solutions have been proposed in the literature (see section 2.2.7 for more details). The second and third problems are related to disambiguation of polysemous words. Word sense disambiguation in information retrieval is a well studied issue on monolingual retrieval (refer to [Sanderson 1996] for a thorough review), however in a cross-language environment, the polysemy problem becomes further complicated.

CLIR is related to machine translation (MT), but the problem of MT is considerably more difficult. MT systems have to pick one translation for each term and they must produce a syntactically correct version of the input text on

the target language. MT is a very difficult problem: it has been studied for many decades and the systems available today are still quite inaccurate.

2.2.3 Brief History

The first investigation on CLIR was done by Salton [Salton 1970], who showed that CLIR systems could perform nearly as well as monolingual systems, given a carefully constructed thesaurus. The first workshop on CLIR was held over 25 years later, in conjunction with SIGIR'96. That gap was due to the lack of available digital resources for searching. Only with the dissemination of information accomplished by the Internet did people realise the importance of CLIR and it was recognised as an independent and multidisciplinary area of research, which combines IR and natural language processing (NLP). At present, CLIR is discussed at conferences on digital libraries, information retrieval, machine translation, and computational linguistics.

2.2.4 Areas of Application

There are several repositories of multilingual information that go beyond national boundaries and the use of a single common language may not be possible or acceptable. CLIR systems can facilitate the IR process in the following areas:

- **Web Search:** the Internet is by far the greatest repository of multilingual information available. The first Web sites were dedicated to the provision of information in English, as the users had mainly academic backgrounds and sufficient English language skills to formulate meaningful queries and understand the results. This, however, is not the current situation. The Web is now used for commerce, entertainment, news, tourism, banking, etc. In addition, the number of non-English speaking users is growing rapidly, there is a prediction that by 2005, they will represent 78% of all users [Peters 2000]. These facts require that the search engines provide Cross-Linguistic facilities.

Solutions for CLIR on Web search are proposed by Braschler [Braschler 1998], Fluhr [Fluhr 1997], Wechsler, Sheridan and Schäuble [Wechsler 1997], and Erbach, Neumann and Uszkoreit [Erbach 1997].

- **Digital libraries:** Recently, there has been a significant growth of interest in the construction of digital libraries throughout the world. They too hold information in many languages, and because they are far more structured and stable than the Web, they make a very attractive area of application to CLIR. Approaches have been suggested by Fluhr [Fluhr 1999], Oard [Oard 1997], and Peters and Picchi [Peters 1997] amongst others.
- **Common Markets** (European Union, NAFTA, MercoSul): The EU has shown a great interest in CLIR, funding several projects on the subject, the most important was perhaps EMIR [EMIR-Consortium 1994]. EMIR was one of the first general CLIR systems to be implemented and evaluated.
- **Multilingual Countries:** In countries where there is more than one official language (e.g. Canada, Switzerland), official information is published in all languages, providing a very good source of corpora (e.g. Canadian parliament debates in French and English; Swiss newspapers in German, French and Italian). Some research that fits within this category has been carried out by Sheridan, Braschler and Schäuble [Sheridan 1997], Landauer and Littman [Landauer 1990], and Davis [Davis 1998].
- **Multinational Companies:** The Intranets of many large organisations also increasingly contain multilingual information. Some companies doing CLIR research are: Xerox [Renders 2003], IBM [Franz 2002] and Microsoft [Gao 2000].

2.2.5 The User of CLIR Systems

There are several types of user that can benefit from CLIR systems:

- A user who has some reading capabilities in a second language but is not confident enough to express his information need in that language. Thus, he

can write the query in his language and the system retrieves the documents in the other language. This scenario is the one which is most considered in CLIR literature.

- A multilingual user who is searching a multilingual collection can be assisted by a CLIR system by entering one single query in a given language and retrieving documents in several languages, avoiding the effort of translating the query in all languages he is proficient in.
- A monolingual user who has no knowledge of a foreign language can benefit from CLIR if he has access to some translation resource. The CLIR system can aid this user by returning highly relevant documents, significantly reducing the number of documents to be translated.

2.2.6 Evaluation of CLIR Systems

Evaluation of CLIR systems is done using the same measurements of traditional monolingual IR, recall and precision. However, the average performance tends to be lower, around 80% of monolingual IR, according to Peters and Sheridan [Peters 2000].

The most important evaluation initiatives are TREC [TREC], CLEF [CLEF] and NTCIR [NTCIR]. TREC has had a special track on CLIR since 1997. CLEF is held in conjunction with the European Conference on Digital Libraries (ECDL), and it is now on its fourth edition. The main difference between them is that CLEF targets European languages, TREC aims at non-European languages and NTCIR deals with Chinese, Japanese and Korean. Aspects of CLIR systems evaluation have been discussed by Braschler, Peters and Schäuble [Braschler 1999], and Braschler [Braschler 2000].

2.2.7 Approaches to CLIR

During the last thirty years four main solutions have been proposed to deal with CLIR: machine translation, thesauri, dictionaries and corpus-based techniques. The next subsections describe each approach.

2.2.7.1 Machine Translation

Machine Translation (MT) systems present the most straightforward approach to CLIR. The first idea that comes to one's mind when thinking about CLIR is to use a fully automatic machine translation (MT) system to bring the query into the same language of the documents. The problem is that queries are usually very short, so they do not provide the MT system with enough context for a high-quality translation. Furthermore, if the MT systems do not have sufficient domain knowledge, they can make errors translating concepts. That can prevent relevant documents from being found and also allows irrelevant documents to be retrieved.

An alternative to this problem would be translating the documents rather than the queries. Because the documents are typically much longer than queries, a MT system would have considerably more contextual information on which to base semantic choices. However, the efficiency of available MT becomes an issue when a large number of documents have to be translated and every time a new language is introduced to the system the whole collection would have to be translated to that language. Nonetheless this approach is growing in popularity. Several recent experiments adopt this strategy; the preferred MT system is the SYSTRAN Software [SYSTRAN]. Gachot, Lange and Yang [Gachot 1998] describe the implementation of the SYSTRAN NLP Browser, a multilingual system. McNamee and Mayfield [McNamee 2000] combine SYSTRAN and n-gram indexing software to achieve better performance than a corpus-based approach. Figuerola [Figuerola 2000] tried several MT systems before choosing SYSTRAN for their Spanish/English experiment finding that the cross-lingual version of the

system performed about 15% worse than the monolingual version. Savoy [Savoy 2003] combined translations provided by 5 freely available MT systems achieving good results on CLEF corpora.

2.2.7.2 Thesauri

A thesaurus is a tool that organises terminology to encode domain knowledge for use by an application; a multilingual thesaurus organises terminology for more than one language. Thesauri typically contain information such as the hierarchical structure among concepts (i.e. broader term, narrower term) and association relationships (i.e. synonyms, antonyms, related terms). In “controlled vocabulary” systems, a specialist (or a computer program) assigns each document in the collection a list of terms from the thesaurus that best describe that document.

CLIR using thesauri works as follows: the user enters the query terms, which are looked up in the thesaurus and replaced by their counterparts in the target language, then a process of monolingual retrieval takes place. Query expansion can be easily performed by using synonyms, which are provided by the thesaurus.

This approach usually achieves good results as thesauri can minimise the problem of polysemous words. However, it has a major disadvantage: thesauri are expensive to build and costly to maintain because they currently need human intervention to recognise and organise the information.

Experiments using thesauri were done by Salton [Salton 1970], Gey and Jiang [Gey 1999], Ferber [Ferber 1997], and Gilarranz, Gonzalo and Verdejo [Gilarranz 1997].

2.2.7.3 Dictionaries

The basic idea of CLIR using dictionaries is to take each query term and translate it using a machine-readable dictionary (MRD). The main motivation behind this approach is that MRDs are becoming increasingly available, both commercially and on-line, for more and more language pairs. In this approach, each word in the query is replaced by all possible translations in the target language (which is a form of query expansion upon translation). The major problem with this approach stems from the fact that MRDs are developed to be explored by humans, requiring some pre-processing before they can be used for CLIR. The following difficulties may occur:

- missing word forms (variant forms of a word may be missing, e.g. words like “women” which is in its plural form);
- spelling norms (color, colour);
- spelling conventions (fallout, fall-out, fall out);
- coverage (technical words are often missing);
- absence of multiword terms (fast food \neq rápido alimento)
- proper names often need to be translated (London, Londres, Londra).

In addition, the dictionary approach suffers significantly from the ambiguity problem. If a polysemous word is present in the query and all possible translations are used, then there will be a large set of target search terms and many of them will be erroneous, resulting in poor precision. Disambiguation techniques for the dictionary approach have been proposed by Adriani [Adriani 2000], Hiemstra and de Jong [Hiemstra 1999], and Hull [Hull 1997].

Some of the most important experiments using dictionaries have been reported by: Hull and Grefenstette [Hull & Grefenstette 1996] who observed that automatic MRD translation can lead to a drop in effectiveness from 40% to 60%; Ballesteros [Ballesteros 2000], and Gollins [Gollins 2000], who tackle the problem of absence of dictionary for a given language pair, using a method called

transitive translation; Ballesteros and Croft [Ballesteros 1997] who showed that translating multi-word phrases can increase CLIR system performance by up to 25%; and Pirkola et al.[Pirkola 2001] who dealt with several of the problems within this approach, including untranslatable words, phrase identification, and ambiguity by using structured queries.

2.2.7.4 Corpus-based Techniques

Corpus-based approaches are strongly motivated by the problems with the thesaurus approach. The aim is to analyse multilingual collections of texts and automatically extract the information needed to construct domain-specific translation techniques. The information about the relationship between terms is obtained from observed statistics of term usage, so thesauri can be automatically constructed.

The collections analysed may consist of parallel (translation equivalent) or comparable (domain specific) sets of documents. The technique known as **document alignment** is widely used on comparable corpora. The alignment process associates documents that cover similar stories. They do it by comparing indicators such as: common proper nouns, common numbers and dates. Let us consider a bilingual collection that contains documents in English and German, the user enters the query in German, the query is run against the German documents, a list of matches is obtained and the matches are replaced by their counterparts in English, if available.

The problem with this approach is the lack of resources: there are few parallel or even comparable multilingual corpus available. For this reason most of the experiments done using this technique were carried out on small and specific collections. Consequently, the translation resources constructed by one experiment would not be suitable for application in any other collections. Some approaches have been proposed to overcome this scarcity problem; Nie and Simard [Nie 2002], and McNamee and Mayfield [McNamee 2003] used a corpus

of parallel pages mined from the web to train translation resources; Lavrenko et al. [Lavrenko 2002] and Orengo and Huyck [Moreira Orengo 2003] used SYSTRAN to translate a sample of documents in order to generate a parallel corpus (see section 4.2).

Peters and Picchi [Peters 1996] propose a method for using comparable corpora, allied to a bilingual dictionary for English/Italian retrieval. Comparable corpora have also been used by Sheridan and Ballerini [Sheridan 1996] who developed the multilingual similarity thesaurus using aligned news stories from the Swiss news agency.

A comparison study done by Yang [Yang 1997] concluded that corpus-based approaches outperform dictionary-based methods. McNamee and Mayfield [McNamee 2002] reached the same conclusion when comparing machine translation systems, bilingual dictionaries, and a large parallel corpus. In their experiments, when a single method was used, best performance was achieved by the corpus approach.

2.2.7.5 Current State of the Art

Strategies for CLIR grow in and out of popularity throughout the years. Recently, MT has been extensively used achieving good results. The preferred MT system is SYSTRAN. Corpus-based techniques have been employed in fewer experiments, however there is a growing trend in using a combination of corpus-based approach with a dictionary or MT in order to improve query expansion. There is a growing interest in stemming and compounding for languages other than English.

Systems that achieve the best results do so by combining several techniques, such as translation resources, stemming (or compounding), more elaborate weighting schemes, query expansion and relevance feedback [Chen 2003;

McNamee 2003; Savoy 2003]. The performance of recently implemented CLIR systems achieves 80 to 95% of the performance of monolingual systems.

2.3 Latent Semantic Indexing

Latent Semantic Indexing (LSI) is a statistical technique developed by Deerwester, Dumais, Furnas, Landauer and Harshman. [Deerwester 1990]. It implements a vector-space model (section 2.1.1.2) and is a corpus-based technique (section 2.2.7.4). This section describes the method and how it can be applied to IR. For a more thorough explanation, please refer to [Deerwester 1990].

2.3.1 Method

The aim is to perform retrieval in terms of conceptual content, rather than term-matching. As the same concept can be expressed in different words (synonymy) and a word can have several meanings (polysemy), individual words provide unreliable evidence about the meaning of the document.

LSI tries to overcome the problems of term matching by treating this unreliability in term usage as a statistical problem. The assumption is that there is some underlying latent semantic structure in the data that is hidden to some extent by the uncertainty in word choice. In order to estimate this latent structure, LSI uses Singular Value Decomposition (SVD) [Forsythe 1977].

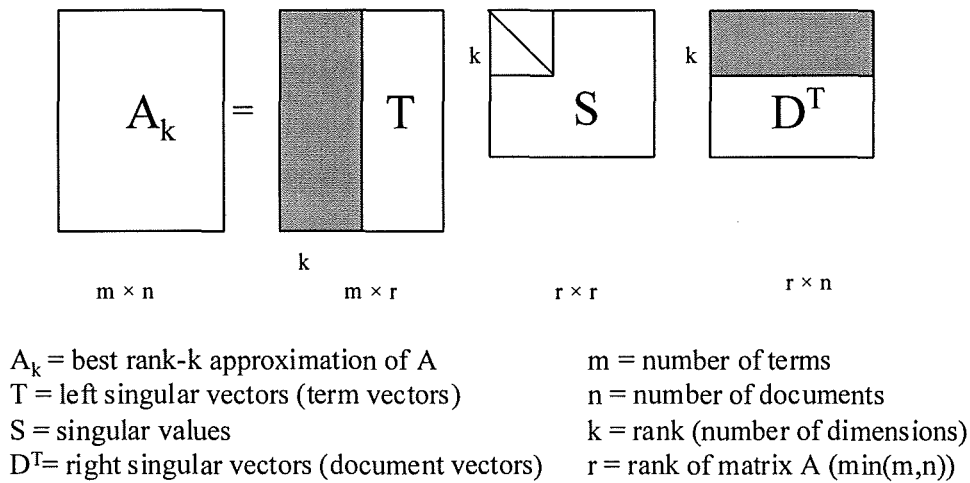
This approach makes use of term co-occurrences to predict which terms are “really” implied by a query. The assumption behind it is that words are not independent from each other and words that frequently co-occur are similar.

LSI uses a term-by-document matrix. The elements of this matrix are the frequencies of occurrence (or weights) of each term in each document. SVD is then applied over this matrix to reduce the number of dimensions. The k most important dimensions are kept. As a result, a semantic space is constructed in

which similar terms and documents are placed close together. The dimensions may be interpreted as artificial concepts representing common meaning extracted from several words and documents. Terms and documents are represented by vectors containing the weights of association with each of those concepts. The SVD of a sparse matrix is given by:

$$A = TSD^T \quad (8)$$

Where T and D have orthonormal columns and S is the diagonal matrix of singular values. The columns of T and D contain the left and right singular vectors, respectively. The $m \times n$ matrix A_k , which is constructed from the k -largest



single triplets of A , is the closest rank- k matrix to A (see Figure 2.6).

Figure 2.6– Singular Value Decomposition

It is possible to reconstruct the original term by document matrix from its factor weights with reasonable accuracy. However, it is not advisable to reconstruct it with perfect accuracy, as the original matrix contains noise, which is eliminated through dimension reduction.

After deriving the semantic space with an initial sample of documents, new documents can be added (folded in), by placing them at the average of their corresponding terms. Because terms and documents are represented on the same semantic space, three types of comparisons are possible: comparing two terms,

comparing two documents, and comparing a term and a document. In order to obtain the similarity between two terms i and j it is necessary to calculate the cosine between the i th and j th rows of the matrix TS . Likewise, to obtain the similarity between two documents i and j it is necessary to calculate the cosine between the i th and j th rows of the matrix DS . Term-document comparisons are a little different, the similarity between term i and document j is obtained by taking the cosine between the i th row of the matrix $US^{1/2}$ and the j th row of the matrix $VS^{1/2}$. This difference is because it is not possible to make a single configuration of points in space that allow both between and within comparisons. Queries are treated as pseudo-documents and placed at the weighted sum of its component term vectors. The representation for the queries is calculated by adding the vector elements for each dimension for all terms and dividing the result by the corresponding singular value (see next section for an example).

2.3.2 Example

Table 2.1 contains titles of documents from the CISI collection. For this example they are the corpus to be indexed by LSI.

d1	Systems Analysis of a University Library; final report and research project
d2	A Library Management Game: a report on a research project
d3	Participative Management in Relation to Library Effectiveness
d4	Case Studies in Systems Analysis in a University Library
d5	The Documentation of Documentation
d6	Computer Handling of Chemical Structure Information
d7	An Efficient Design for Chemical Structure Search. I. The Screens
d8	The IDC System for Chemical Documentation
d9	The GREMAS System, an Integral Part of the IDC System for Chemical Documentation
d10	A Chemical Search System for a Small Computer

Table 2.1 – Sample collection of documents

The first step is to generate a term-by-document matrix of the texts. The cells of that table contain the frequency of occurrence of each term in each document. Table 2.2 presents the term-by-document Matrix *A*, the terms were stemmed, and stopwords were discarded. In addition, LSI only indexes terms appearing in more than one document¹:

Terms	Documents									
	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
(t1) analysis	1	0	0	1	0	0	0	0	0	0
(t2) chemical	0	0	0	0	0	1	1	1	1	1
(t3) computer	0	0	0	0	0	1	0	0	0	1
(t4) documentation	0	0	0	0	2	0	0	1	1	0
(t5) idc	0	0	0	0	0	0	0	1	1	0
(t6) library	1	1	1	1	0	0	0	0	0	0
(t7) management	0	1	1	0	0	0	0	0	0	0
(t8) project	1	1	0	0	0	0	0	0	0	0
(t9) report	1	1	0	0	0	0	0	0	0	0
(t10) research	1	1	0	0	0	0	0	0	0	0
(t11) search	0	0	0	0	0	1	1	0	0	1
(t12) structure	0	0	0	0	0	1	1	0	0	0
(t13) system	1	0	0	1	0	0	0	1	2	1
(t14) university	1	0	0	1	0	0	0	0	0	0

Table 2.2 – Term-by-document matrix *A*

The next step is to calculate the SVD for matrix *A*. The decomposition will generate 3 matrices: a matrix of left singular vectors (or term vectors), a diagonal matrix of singular values, and a matrix of right singular vectors (or document vectors). Those are shown in Figure 2.7.

¹ The procedure aims at reducing the number of index terms. However, Efthimiadis' experiments on query expansion [Efthimiadis 2000] give support to this procedure as he concluded that words occurring in only one document contribute to poor retrieval performance.

14 term vectors by 10 dimensions (Matrix T)

(t1)	-0.1831	0.2533	-0.0090	0.3022	-0.3899	-0.0499	0.0906	-0.1360	0.1724	-0.2447
(t2)	-0.4127	-0.3723	-0.2771	-0.0716	0.1657	0.0208	0.2132	-0.0526	0.4469	-0.0194
(t3)	-0.1340	-0.1535	-0.3396	0.0009	-0.0646	-0.0104	-0.6909	-0.5750	-0.1756	-0.0000
(t4)	-0.3142	-0.2378	0.5223	-0.5582	-0.4923	-0.0281	-0.1256	0.0359	-0.0059	0.0000
(t5)	-0.2369	-0.1515	0.2263	-0.0040	0.3443	0.0349	0.3218	-0.5240	0.1960	0.0194
(t6)	-0.2505	0.4611	-0.1007	-0.1049	-0.0361	-0.5377	0.0401	-0.0658	0.0849	-0.4070
(t7)	-0.0674	0.2078	-0.0917	-0.4071	0.3538	-0.4878	-0.0505	0.0702	-0.0875	0.4070
(t8)	-0.1536	0.3168	-0.0875	-0.2053	0.0658	0.3937	-0.0120	0.0031	-0.0017	-0.2887
(t9)	-0.1536	0.3168	-0.0875	-0.2053	0.0658	0.3937	-0.0120	0.0031	-0.0017	0.3148
(t10)	-0.1536	0.3168	-0.0875	-0.2053	0.0658	0.3937	-0.0120	0.0031	-0.0017	-0.0262
(t11)	-0.1758	-0.2208	-0.5034	-0.0677	-0.1786	-0.0141	-0.1086	0.4713	0.2509	0.0194
(t12)	-0.0919	-0.1486	-0.3757	-0.1369	-0.2555	-0.0134	0.5493	-0.1505	-0.6479	0.0000
(t13)	-0.6422	-0.0483	0.2182	0.4172	0.2583	0.0023	-0.1525	0.3157	-0.4170	0.0000
(t14)	-0.1831	0.2533	-0.0090	0.3022	-0.3899	-0.0499	0.0906	-0.1360	0.1724	0.6516

Square matrix of Singular Values (Matrix S)

4.0332	0	0	0	0	0	0	0	0	0
0	3.3185	0	0	0	0	0	0	0	0
0	0	2.6570	0	0	0	0	0	0	0
0	0	0	2.0072	0	0	0	0	0	0
0	0	0	0	1.5342	0	0	0	0	0
0	0	0	0	0	1.3352	0	0	0	0
0	0	0	0	0	0	1.0597	0	0	0
0	0	0	0	0	0	0	0.5063	0	0
0	0	0	0	0	0	0	0	0.3421	0
0	0	0	0	0	0	0	0	0	0.0000

10 Document Vectors by 10 dimensions (Matrix D)

d1	-0.4264	0.5635	-0.0613	0.1498	-0.2347	0.4090	0.0310	-0.0253	0.0219	-0.5000
d2	-0.1931	0.4880	-0.1712	-0.5620	0.3357	0.1167	-0.0437	0.0269	-0.0225	0.5000
d3	-0.0788	0.2015	-0.0724	-0.2550	0.2070	-0.7680	-0.0097	0.0087	-0.0074	-0.5000
d4	-0.3121	0.2770	0.0375	0.4567	-0.3634	-0.4756	0.0650	-0.0435	0.0370	0.5000
d5	-0.1558	-0.1433	0.3931	-0.5563	-0.6418	-0.0420	-0.2371	0.1417	-0.0345	0
d6	-0.2019	-0.2697	-0.5630	-0.1371	-0.2170	-0.0128	-0.0350	-0.6060	-0.3675	-0.0000
d7	-0.1687	-0.2235	-0.4351	-0.1376	-0.1749	-0.0050	0.6171	0.5298	0.1459	-0.0000
d8	-0.3982	-0.2440	0.2596	-0.1079	0.1800	0.0224	0.2425	-0.4444	0.6432	-0.0000
d9	-0.5574	-0.2586	0.3418	0.0999	0.3483	0.0242	0.0986	0.1791	-0.5761	-0.0000
d10	-0.3383	-0.2395	-0.3394	0.1389	0.1179	-0.0010	-0.6972	0.3148	0.3074	0.0000

Figure 2.7 – SVD Example

The matrices in Figure 2.7 show the complete SVD for matrix A . The singular values are organised in decreasing order and quantify the importance of the dimension they represent. The reason LSI applies SVD is to discard the less important dimensions. In this example only the two most important dimensions will be kept (for visualisation purposes). Figure 2.8 shows the vectors after the dimension reduction.

T_2	S_2	D_2^T
-0.1831	0.2533	4.0332
-0.4127	-0.3723	0
-0.1340	-0.1535	3.3185
-0.3142	-0.2378	0.5635
-0.2369	-0.1515	0.4880
-0.2505	0.4611	0.2015
-0.0674	0.2078	0.2770
-0.1536	0.3168	-0.1433
-0.1536	0.3168	-0.2697
-0.1536	0.3168	-0.2235
-0.1758	-0.2208	-0.2440
-0.0919	-0.1486	-0.2586
-0.6422	-0.0483	-0.2395
-0.1831	0.2533	

Figure 2.8 – Dimension Reduction

Figure 2.9 shows the representation of the terms and documents in the semantic space. Terms are represented by circles and documents are represented by crosses. Note that documents 1 to 4 and documents 5-10 are placed close together as they share words. Terms that have high co-occurrence (t_8, t_9, t_{10}) are treated as synonyms and placed at the same point.

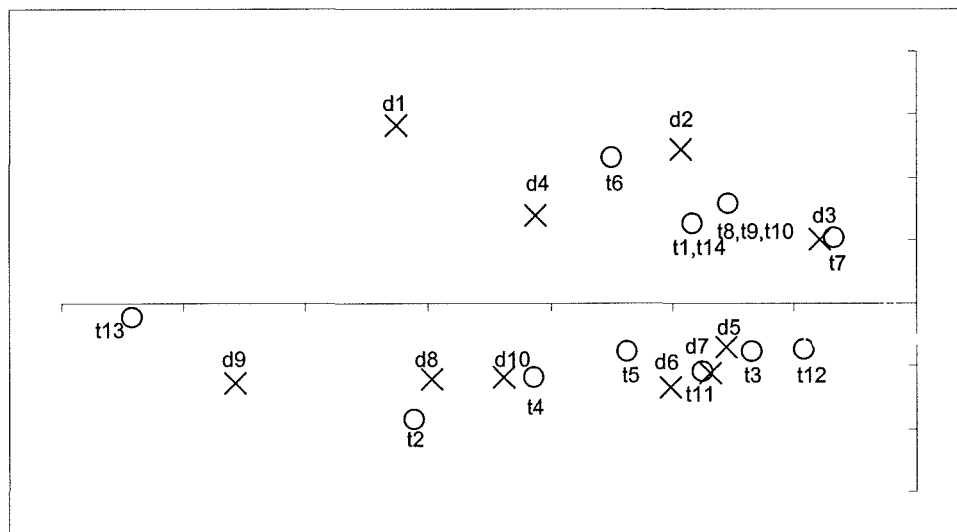


Figure 2.9 – Plot of terms and documents using two dimensions

The results of the cosine for term t_8 (project) with all other terms is shown on Table 2.3. The calculations are done using rows of the matrix $T_2 S_2$. Term t_8 was assigned identical vectors to t_9 and t_{10} , therefore its cosine with those term vectors is equal to 1. The most dissimilar term to term t_8 is t_{11} (search), as they do not co-occur in any documents.

Terms	Cosine with t8
(t1) analysis	0.98
(t2) chemical	-0.11
(t3) computer	-0.22
(t4) documentation	-0.02
(t5) idc	0.05
(t6) library	1.00
(t7) management	0.99
(t8) project	1.00
(t9) report	1.00
(t10) research	1.00
(t11) search	-0.27
(t12) structure	-0.38
(t13) system	0.45
(t14) university	0.98

Table 2.3 – Term-term comparisons

In the same way, it is possible to calculate similarities between documents. Table 2.4 shows the similarity between document *d1* and all other documents. That similarity will reflect the extent to which the documents share the same terms. This is done by taking the cosine between rows of the matrix D_2S_2 .

Documents	Cosine with d1
d1	1.00
d2	0.96
d3	0.96
d4	0.98
d5	0.10
d6	-0.09
d7	-0.08
d8	0.27
d9	0.37
d10	0.21

Table 2.4 – Document-document comparisons

Now, let us consider the situation in which a user poses the query “library system”. The first step is to assign a vector representation for the query. The first element of the vector is calculated by adding the first elements for the two term vectors and dividing the result by the singular value corresponding to the first dimension. The second element of the vector is calculated in a similar way. Thus the vector for the query (or pseudo-document) is: $Q = (-0.2213, 0.1244)$. The next

step is to calculate the cosine between this query and all documents. The result is then sorted in decreasing order of similarity. This is shown in Table 2.5.

Documents	Cosine with Q
d4	0.99
d1	0.95
d2	0.82
d3	0.82
d9	0.64
d8	0.56
d10	0.51
d5	0.40
d7	0.23
d6	0.22

Table 2.5 – Query-document comparison

It is possible to approximate the original matrix A (Table 2.2) by multiplying T , S and D^T . Table 2.6 shows the approximation of A using 2 dimensions. The more dimensions kept the closer to the original matrix A .

$A_2 =$	0.7885	0.5527	0.2276	0.4634	-0.0054	-0.0776	-0.0633	0.0890	0.1943	0.0486
	0.0137	-0.2814	-0.1177	0.1774	0.4364	0.6693	0.5569	0.9643	1.2473	0.8591
	-0.0566	-0.1442	-0.0600	0.0276	0.1572	0.2464	0.2050	0.3394	0.4328	0.3048
	0.0956	-0.1405	-0.0592	0.1769	0.3106	0.4688	0.3902	0.6972	0.9105	0.6178
	0.1243	-0.0607	-0.0260	0.1591	0.2209	0.3285	0.2735	0.5032	0.6626	0.4437
	1.2930	0.9417	0.3880	0.7393	-0.0619	-0.2087	-0.1715	0.0290	0.1676	-0.0245
	0.5045	0.3890	0.1604	0.2759	-0.0565	-0.1311	-0.1082	-0.0600	-0.0267	-0.0731
	0.8566	0.6327	0.2607	0.4846	-0.0542	-0.1585	-0.1305	-0.0099	0.0734	-0.0422
	0.8566	0.6327	0.2607	0.4846	-0.0542	-0.1585	-0.1305	-0.0099	0.0734	-0.0422
	0.8566	0.6327	0.2607	0.4846	-0.0542	-0.1585	-0.1305	-0.0099	0.0734	-0.0422
	-0.1106	-0.2207	-0.0918	0.0183	0.2155	0.3408	0.2834	0.4611	0.5846	0.4154
	-0.1199	-0.1691	-0.0702	-0.0210	0.1284	0.2079	0.1728	0.2679	0.3341	0.2435
	1.0141	0.4219	0.1719	0.7640	0.4265	0.5661	0.4727	1.0704	1.4851	0.9146
	0.7885	0.5527	0.2276	0.4634	-0.0054	-0.0776	-0.0633	0.0890	0.1943	0.0486

Table 2.6 – Matrix A_2 – Approximation of matrix A using 2 dimensions

2.3.3 Other applications

LSI has been applied to other tasks, besides IR. This section outlines some of the most important:

- **Information Filtering:** Foltz [Foltz 1990] proposed using LSI for Information Filtering, which is a similar task to IR with the difference that the queries are static and the document database is constantly changing. The paper used LSI to predict whether new articles would be judged interesting based on previous judgements. The results have shown improvements in precision of 13% over keyword matching. In an experiment done by Foltz and Dumais [Foltz 1992], LSI outperformed three other methods in an information filtering experiment.
- **Assigning Manuscripts to Reviewers:** Dumais and Nielsen [Dumais 1992] applied LSI to the process of assigning papers to reviewers in a conference. This assignment was done based on the reviewer's interests. The system achieved a reasonable performance, scoring slightly poorer than human experts.
- **Marking Student Essays:** Redher et al. [Redher 1998] show how LSI can be used to mark student essays. The system ranks the essay according to its cosine similarity to the model. Lemaire and Dessus [Lemaire 2001] performed several tests and conclude that LSI's performance is comparable to humans'.
- **Text Classification:** Zelikowitz and Hirsh [Zelikowitz 2001] used LSI for text classification, and achieved superior results to a system using a machine learning method.

2.3.4 Discussion of LSI

There are some aspects on the use of LSI for IR that remain unclear. The only currently known way to determine the number of dimensions is through trial and error as there is no automatic way of establishing the optimal number. LSI research suggests a number which is large enough to incorporate important concepts and small enough exclude noise. In previous experiments, this number was typically between 100 and 300 (see section 4.4.2. for experiments with different number of dimensions).

One of the drawbacks of the method is that SVD is computationally expensive and it may take several hours to be computed, for a reasonably sized collection. In addition, the vectors generated are dense, and that delays the calculation of the cosine measurement during the processing of queries.

2.4 CLIR Using Latent Semantic Indexing

The main goal of using LSI for CLIR is to provide a method for matching text segments in one language with text segments of similar meaning in another language without needing to translate either; LSI does this by creating a language-independent representation of the words. Words are given an abstract description that does not depend on the original language.

LSI is initially applied to a matrix of terms by documents (see Figure 2.10). Therefore, the first step is to build such a matrix based on a set of dual-language documents². The matrix contains the number of occurrences (or weights) of each term in each document. In an ideal situation the pattern of occurrence of a term in language A should be identical to the pattern of occurrence of its match in language B. However, in the non-ideal real world, those patterns are not identical. Nonetheless, the similarity between a term and its translation is still quite high (see section 4.4.5 for experiments).

This matrix is then factorised by SVD. After deriving the semantic space with an initial sample of documents, new documents can be added at the average of their corresponding terms.

² Dual-language documents are composed of the document in the original language together with its translation in another language.

		D ₁	D ₂	D ₃	D ₄	D ₅	D ₆	D ₇	...	D _n
A	água	3	0	4	0	0	1	1	...	0
	casa	0	0	2	1	1	1	0	...	0
	papel	0	0	0	0	0	0	3	...	1
	porta	0	0	1	1	0	1	1	...	0
	...									
B	door	0	0	1	1	0	1	1	...	0
	house	0	0	2	1	1	1	0	...	0
	paper	0	0	0	0	0	0	3	...	1
	water	3	0	4	0	0	1	1	...	0
	...									

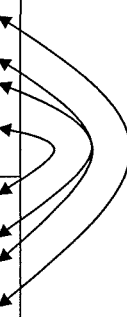


Figure 2.10 – Bilingual term by document matrix

SVD causes synonyms to be represented by similar vectors (since they would have many co-occurrences), which allows relevant documents to be retrieved even if they do not share any terms with the query. This is what makes LSI suitable for CLIR, given that a term in one language will be treated as a synonym to its match in the other language. The main advantages of using LSI for CLIR are:

- There is no translation. All terms and documents are transformed to a language-independent representation.
- New languages can be added easily.
- There is no need for expensive resources such as dictionaries, thesauri or machine translation systems.

The first experiment on CLIR using LSI was carried out by Landauer and Littman [Landauer 1990]. The bilingual corpus used was composed of 2482 paragraphs from the Proceedings of the Canadian Parliament (French/English). They took a sample of 900 dual-language paragraphs to derive the semantic space. Then they folded-in the remaining 1582 English-only documents. Next, the French-only paragraphs were used as queries and 92% of the times the top retrieved document was the corresponding English document. It is important to mention that this test was just a mate-finding experiment and not a standard IR evaluation. Redher et al. [Redher 1997] used LSI to participate in the TREC cross-language track and the performance ranged from 8 to 14% average precision,

which was an average result for the task (i.e. similar to other systems).

Yang [Yang 1997] performed tests with query translation and statistical IR methods. In all tests, LSI's performance was among the best. In addition, the loss in performance between monolingual and bilingual executions was small, about 15%.

2.5 Relevance Feedback

The Relevance Feedback (RF) operation is an automatic process for the modification of search requests based on relevance assessments provided by the user population for previously retrieved documents [Salton 1971a]. The idea behind it is that users are unlikely to produce perfect queries, especially if given just one attempt. The typical process improves the query specification by choosing important terms attached to previously retrieved documents that have been identified as relevant by the user. This will lead to more relevant documents being retrieved. The method, illustrated in Figure 2.11, consists of the last three steps of the following process:

1. The searcher enters a query.
2. The system prepares a ranked list of documents.
3. The searcher judges some of the documents for relevancy.
4. The system reconstructs the original query based on the judgements provided.
5. A new ranked list of documents is prepared. The assumption is that this list will be better than the previous.

RF is an iterative method that can be repeated several times, until the user is satisfied with the result. Salton and Buckley [Salton 1997] report on the advantages of RF:

- It shields users from the details of the query reformulation process.
- It breaks down the search operation into a sequence of small search steps, designed to approach the wanted subject gradually.

- It provides a controlled query alteration process designed to emphasise some terms and de-emphasise others, as required in particular search environments.

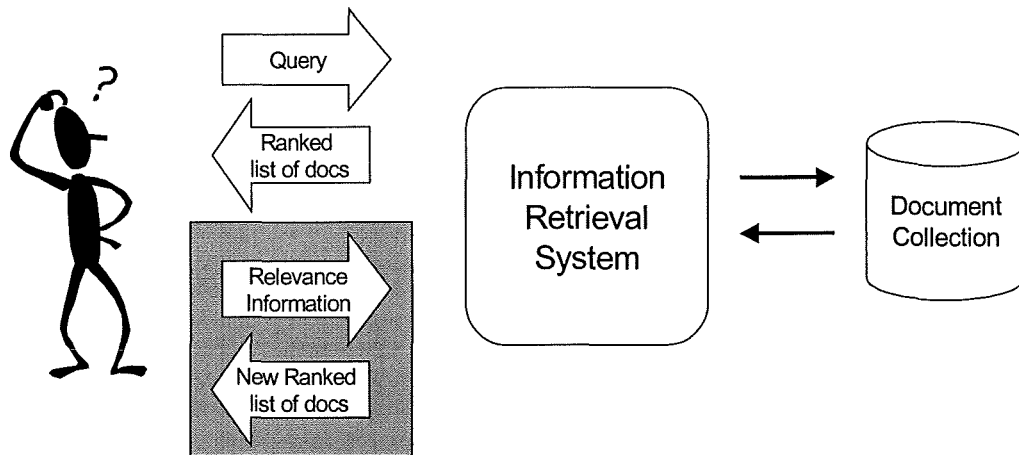


Figure 2.11 Relevance Feedback Process

RF can be performed without user interference, in a technique known as pseudo-relevance feedback (PRF). In PRF, n top ranked documents are assumed relevant and used for the feedback runs. This technique achieves less improvement than original RF, however it has the advantage of being done automatically without any burden to the user.

2.5.1 Brief History

The concept of RF was introduced in the mid-1960s. Some early experiments were done by Rocchio on the SMART Retrieval system [Rocchio 1971], which implements a vector-space model. Since then, the method has been applied to other IR approaches. In 1976, Robertson and Sparck Jones [S. Robertson 1976] have experimented with RF on a probabilistic model; and Dillon and Desper [Dillon 1980] proposed an RF method for a boolean Retrieval System. In addition, several different feedback procedures have been proposed. Some of the best known were developed by Ide [Ide 1971]. More recently, new approaches include Neural Networks [Crestani 2000], Genetic Algorithms [Biron 1995], and Machine Learning [Drucker 2001].

2.5.2 Relevance Feedback and the Different IR Models

This section describes how RF can be performed on different IR models, namely vector-space, Boolean, probabilistic and LSI. RF may involve query expansion or may just reweigh terms present in the initial query reformulation based on relevance information. RF methods that include query expansion are known to perform better than the ones that simply reweigh query terms.

2.5.2.1 Relevance Feedback and the Vector-Space Model

The original RF process was designed to be used with vector queries, which are written without Boolean operators on the form:

$$Q_0 = (q_1, q_2, \dots, q_l) \quad (9)$$

where q_1 represents the weight of term 1 in query Q_0 . A weight of 0 indicates that the term is not present in the query and a weight of 1 indicates that the term has full weight. The RF process will generate a modified version of the query, which can be written as

$$Q' = (q'_1, q'_2, \dots, q'_l) \quad (10)$$

where q'_1 represents the modified term weight for term 1. New terms can be added by modifying an initial weight from 0 to some positive value. The expected result is that the query vector will be moved away from the irrelevant documents and closer to the relevant ones.

The first three methods of RF in the vector-space model are described below [Harman 1992]. The basic idea of these methods is to merge document vectors and the initial query vector.

- Ide Regular $Q' = Q_0 + \sum_{k=1}^{n_1} R_k - \sum_{k=1}^{n_2} S_k \quad (11)$

- Ide dec-hi $Q' = Q_0 + \sum_{k=1}^{n_1} R_k - S_1 \quad (12)$

$$\bullet \text{ Rocchio} \quad Q' = Q_0 + \beta \sum_{k=1}^{n_1} \frac{R_k}{n_1} - \gamma \sum_{k=1}^{n_2} \frac{S_k}{n_2} \quad (13)$$

where, R_k is the vector for the relevant document k

S_k is the vector for the non-relevant document k

n_1 is the number of relevant documents

n_2 is the number of non-relevant documents and

β and γ are parameters that control the contributions of relevant and non-relevant documents.

All three methods keep the initial query as it may contain important information. Rocchio's method provides the constants β and γ to weight the importance of the information contained in relevant and non-relevant documents. Generally, it is found that relevant documents provide more important information than non-relevant ones, therefore β can be made greater than γ . For cases in which only positive feedback is sought, γ can be set to 0. Ide Regular works in a similar way, adding terms found in relevant documents and subtracting terms from non-relevant documents. The difference of the Ide dec-hi method is that only the top non-relevant document is used for negative feedback.

Salton and Buckley [Salton 1997] have evaluated all three methods and concluded that Ide Dec-hi performs slightly better. In further experiments Harman [Harman 1992] recommends using only 20 selected terms from the relevant documents, rather than all the terms.

2.5.2.2 Relevance Feedback and the Boolean Model

Only a few experiments have tested RF for Boolean systems. The feedback process in Boolean systems is more difficult than on the vector model for two reasons: the choice of terms is more crucial, and the model requires the choice of Boolean operators to relate the terms.

Dillon and Desper [Dillon 1980] developed a technique that uses relevance information provided by the user to calculate weights for the terms in the retrieved documents. The terms are ordered by these weights in the construction of the new Boolean query. The weights range from -1 to 1. The reformulated query will have two parts: one with the terms with high weights (good terms) and another with terms with low weights (bad terms). The good terms are divided into two groups. Each term from the first group (the highest ranked) is *or'd* (t_1 *or* t_2 *or* ... *or* t_n). Terms in the second group are *and'd* in pairs (t_1 *and* t_2), each term is paired with every other term in the group. Bad terms (terms with scores approaching -1) are also divided into two groups. Terms from the worse group are used in an expression of the form: *not* t_1 *and* *not* t_2 *and* ... *and* *not* t_n . Terms from the second worse group are *and'd* in pairs in the form *not* (t_1 *and* t_2), for all pairs. Tests on a small collection have shown improvements comparable to RF methods used on vector-space IR systems.

Salton, Fox and Voorhees [Salton 1985] describe a method for Boolean RF composed by two steps: the construction of good term clauses, and the generation of a modified Boolean query using some of the previously chosen clauses. The use of the clause (which can be a single term or a set of terms connected by *and* operators) depends on its relevance weight and its expected postings frequency. The relevance weight measures the degree to which the clause is likely to be useful for retrieving relevant documents. The postings frequency represents the approximate number of documents which the clause is expected to retrieve. The reformulated feedback query consists of a set of clauses connected by *or* operators. The approach has been tested and achieved improvements ranging from 22% (on the CISI collection) to 110% (on the Medlars collection).

2.5.2.3 Relevance Feedback and the Probabilistic Model

The probabilistic model is based on the distribution of query terms over relevant and nonrelevant documents. Robertson and Sparck Jones [S. Robertson 1976] developed a formula for term weighting based on such distribution.

$$w_{ij} = \log_2 \frac{\frac{r}{R-r}}{\frac{n-r}{N-n-R+r}} \quad (14)$$

where: w_{ij} = the weight for term i in query j

N = the number of documents in the collection

R = the number of relevant documents for the query j

n = the number of documents having the term i

r = the number of relevant documents having the term i

This formula assumes that all the relevant documents are known, which is not the case for real world applications. Therefore some estimates are required. Tests have been performed using the Cranfield 1400 collection and achieved significant improvement compared to keyword matching.

According to Baeza-Yates and Ribeiro-Neto [Baeza-Yates 1999] the main advantages of this method is that the feedback process is directly related to the derivation of new weights for the query terms. The disadvantages include: term-document weights are not taken into account for the feedback loop; weights calculated in previous query formulations are also disregarded and query expansion is not used, only terms present in the initial query formulation are reweighted. Croft [Croft 1983] proposed an extension to this weighting scheme which addresses some of the mentioned disadvantages. He suggested effective initial search methods and takes into consideration the term-document weights.

2.5.2.4 Relevance Feedback and Latent Semantic Indexing

RF can be easily performed in a LSI system. The method proposed by Dumais [Dumais 1991] consists of replacing the original query with the vector sum of the documents the user has judged relevant.

Her experiments compared the performance of RF using three different methods: (i) using the highest ranked relevant document; (ii) using the weighted average (or centroid) of the first three relevant documents and (iii) using the centroid of all relevant documents. Experiments have been carried out using 5 small test collections (ADI, CISI, MED, CRAN and Time), the results show improvements of 67% on average. The best feedback method was replacing the query with the centroid of all relevant documents. This approach is not realistic, as it is not possible to know all the relevant documents in advance. This would require the user to analyse all documents in the collection for relevance. In addition, because of the evaluation strategy used, part of the improvement observed can be attributed to the re-ranking of documents the user has judged relevant (for a discussion on the evaluation of RF please see section 2.5.3).

The study also observed a big variation on performance improvement for different collections. The MED collection had a very small gain while the CISI collection had its average precision almost trebled. In general, the higher the initial performance, the smaller the improvement obtained.

An important point is that using documents as queries does not impose any extra costs in terms of system computation. Since both terms and documents are represented as vectors in the same semantic space, the formation of queries as a combination of terms or documents is straightforward.

All these approaches only make use of positive feedback. There is no attempt to move the query away from the non-relevant documents by subtracting the vectors of the documents the user has judged non-relevant.

2.5.2.5 Other Methods

RF is still a popular research topic and new approaches are constantly being proposed. Xu and Croft [Xu 1996] proposed a method called Local Context Analysis (LCA) that combines PRF and global analysis. Global analysis methods extract concepts from text, based on word co-occurrences using statistical techniques. They report an improvement of 23% on retrieval performance on the TREC4 collection against an 11% improvement produced by PRF.

Drucker et al. [Drucker 2001] apply Support Vector Machines (SVM) to the problem. The idea is to use SVMs to separate the hyperplane³ containing relevant documents (or vectors) from the hyperplane containing nonrelevant ones. Tests have been performed using the Reuters corpus and ten feedback iterations. SVM's performance has been comparable to the Ide dec-hi algorithm, and superior in cases in which the query topic occurred few times across the collection.

Biron and Kraft [Biron 1995] show how genetic algorithms (GAs) can be applied for RF in boolean queries. A GA is initialised with a set of queries and relevant documents, and evolves new queries. The generated queries are evaluated using a fitness function. Experimental results were not presented.

Crestani [Crestani 2000] uses a back-propagation learning algorithm for implementing neural RF. The method consists in training a 3-layer feedforward neural network using queries and documents known relevant to produce a set of terms to be used for feedback. In the spreading phase, query terms are presented and the higher activated m nodes (or terms) are added to the query. The paper also reports on tests carried out using the Cranfield, CACM and NPL collections.

³ A hyperplane is an N -dimensional analogy of a line or a plane, which divides an $N+1$ dimensional space into two. E.g. a in a three-dimensional space, hyperplane is the usual two-dimensional plane.

Results achieved by this method were smaller than the ones achieved by probabilistic RF.

Lam-Adesina and Jones [Lam-Adesina 2001] proposed a method for using terms extracted from document summaries for pseudo relevance feedback. Document summaries are built taking into consideration several aspects: keyword clustering, the position of sentences, the number of title terms contained in the sentence, and the number of query terms in the sentence. Results show an improvement of up to 15% using the TREC collection.

2.5.3 Evaluation

A simple approach to the evaluation of RF would be measuring recall and precision before and after RF and compare both. This is how the fluid relevance feedback technique [Chang 1971] works. However, this is not a good measure of improvement as performance can increase simply because the documents classified as relevant by the user will be ranked higher on the feedback retrieval set, and the irrelevant documents discarded. Williamson [Williamson 1978] suggests that most of the merit for RF strategies is due to the searcher's individual decisions, and not due to the ability of the system in making an additional contribution. This is known as the "ranking effect" and it leads to dramatic improvements masking the real advantage provided by the process.

A measure of effectiveness of a relevance feedback system should reflect the number of newly retrieved relevant documents due to the feedback process. Some methods have been proposed aiming at measuring the real improvement provided by RF. The best known approaches are described below.

- **Full Freezing:** This method, used by the SMART [Salton 1971b] system, freezes the ranks of all documents presented to the user in the feedback iterations and assigns to the first document retrieved by the i^{th} iteration a rank $iN+1$ (where N documents are presented to the user for feedback). This

strategy gives results that are misleadingly low as top ranks are frozen, so no newly retrieved document can achieve a higher rank than documents retrieved in previous iterations.

- **Modified Freezing:** This method, proposed by Ide [Ide 1971], differs from Full Freezing in that non-relevant documents ranked below the last relevant document are not frozen, hence the number of documents frozen at each iteration vary. This technique provides more realistic results than the previous method. However, it can only be used for positive feedback since documents judged as irrelevant do not have their ranks frozen, which allows them to be used again to modify the query.
- **Residual Collection:** This is the most used approach for the evaluation of RF. In this method, the documents that have been judged by the searcher are excluded from the collection and from the relevance assessments. The second ranked list will only contain documents that have not been judged. This method provides an unbiased evaluation of RF, however, the elimination of highly ranked relevant documents that were present on the initial run will lower recall-precision figures. Therefore it is not possible to directly compare the results of this method with results calculated using the fluid technique.

2.5.4 Discussion of Relevance Feedback

Several experiments report on the performance improvement achieved by relevance feedback; Salton and Buckley [Salton 1997] experimented with several RF methods. The improvement achieved (measured using the residual collection approach) ranged between 47% (for the CISI collection) and 160% (for the Cranfield collection). Harman [Harman 1992] reports improvements of 112% for the Cranfield collection when expanding the query with 20 terms chosen from the relevant documents.

Salton and Buckley [Salton 1997] also concluded that some types of collections may benefit more from the RF process. These are collections with short queries; collections with queries that perform relatively poorly in an initial search; and technical collections. Collections with short queries can benefit from the RF process as it will add context, making the query more complete. Collections with queries that have a weak performance on the initial run have more potential for improvement. In technical collections, it is possible that the set of relevant documents for a given query is concentrated in a small area of the document space.

For their CLEF experiments, McNamee and Mayfield [McNamee 2002] report that RF does not always improve performance, and in some cases can even decrease it. They observed this problem especially when the initial queries are longer. However, those findings were based on PRF which normally performs worse than user RF.

Despite the great research interest on RF, very few experiments have been carried out using human searchers. A user-centered investigation has been made by Efthimiadis [Efthimiadis 2000]; he observed 25 searchers querying the INSPEC database. The main findings confirm the effectiveness of RF. The initial search produced on average three highly relevant documents, and the feedback run produced on average nine further highly relevant documents. Another study involving users was done by Spink [Spink 1994]; she performed an experiment with forty users to assess how humans perform query expansion in an interactive environment. She concluded that users are able to select effective terms for query expansion. However, because the study used real user queries, it is not possible to calculate evaluation measures and thus quantify the gain obtained from reformulation.

2.6 Summary of the Chapter

This chapter has presented a literature review on Information Retrieval, Cross-Language Information Retrieval, Latent Semantic Indexing and Relevance Feedback. An introduction on the main points of each subject was given, and pointers to the relevant literature were provided. The concepts presented here will be applied in the experiments reported in the subsequent chapters.

3 A STEMMING ALGORITHM FOR PORTUGUESE

Before carrying out any IR experiments, it is necessary to gather some resources. Because IR research in Portuguese is still in its infancy, there is a lack of linguistic tools available for the research community. Until now the only available stemmer for Portuguese is a translated version of a well known algorithm [Porter 1980] that has been designed for English. The problem in that algorithm is that the morphology of Portuguese has not been taken into account. Therefore the first step towards a Portuguese-English CLIR system was to create a suitable stemmer for Portuguese to be used by the IR experiments described in the subsequent chapters.

This chapter describes the development of a simple and effective suffix-stripping algorithm for Portuguese. The stemmer is evaluated using a method proposed by Paice [Paice 1994]. The results show that it performs significantly better than the Portuguese version of the Porter algorithm, which is the only other Portuguese stemmer publicly available. The algorithm discussed here was initially presented in [Moreira Orengo 2001] and will be referred to as RSLP (Removedor de Sufixos da Língua Portuguesa).

3.1 Introduction

Stemming is the process of conflating the variant forms of a word into a common representation, the stem. This procedure is widely used in IR with the aim of

increasing recall. Several studies evaluate the validity of the stemming process for IR and they have reached contrasting conclusions. Harman [Harman 1991] examined the effects of 3 algorithms on 3 test collections and found no improvements on the retrieval performance since the number of queries with improved performance tended to equal the number with poorer performance. However in [Krovetz 1993], stemming improved retrieval performance by up to 35% on some collections. Finally, after exhaustive analysis Hull [Hull 1996] concludes that “some form of stemming is almost always beneficial”, he found that the overall improvement ranged from 1-3% but, for many individual queries, stemming made a large difference. All those experiments were done on English collections and it seems possible that highly inflected languages such as Portuguese may benefit more from stemming.

The mistakes associated with stemming can be divided in two groups:

- Overstemming: when the string removed was not a suffix, but part of the stem. This can result in the conflation of unrelated words.
- Understemming: when a suffix is not removed. This will cause a failure in conflating related words.

The stemming of English seems to be a largely resolved problem, the Porter Stemmer [Porter 1980] is a simple suffix-stripping algorithm that is based solely on rules, without exception lists or dictionary lookups, however it has proved to be as effective as more elaborated systems. Similar algorithms have been developed for other languages [Honrado 2000; Kraaij 1994; Wechsler 1997]. The challenge was to design a suffix-stripping algorithm for Portuguese that is both simple and effective with the target of improving recall, without decreasing precision.

3.2 The Algorithm

The RSLP algorithm was implemented in C and is composed of 8 steps that need to be executed in a certain order. Figure 3.1 shows the sequence of those steps

Each step has a set of rules, the rules in the steps are examined in sequence, and only one rule in a step can apply. The longest possible suffix is always removed first because of the order of the rules within a step, e.g. the plural suffix *-es* should be tested before the suffix *-s*.

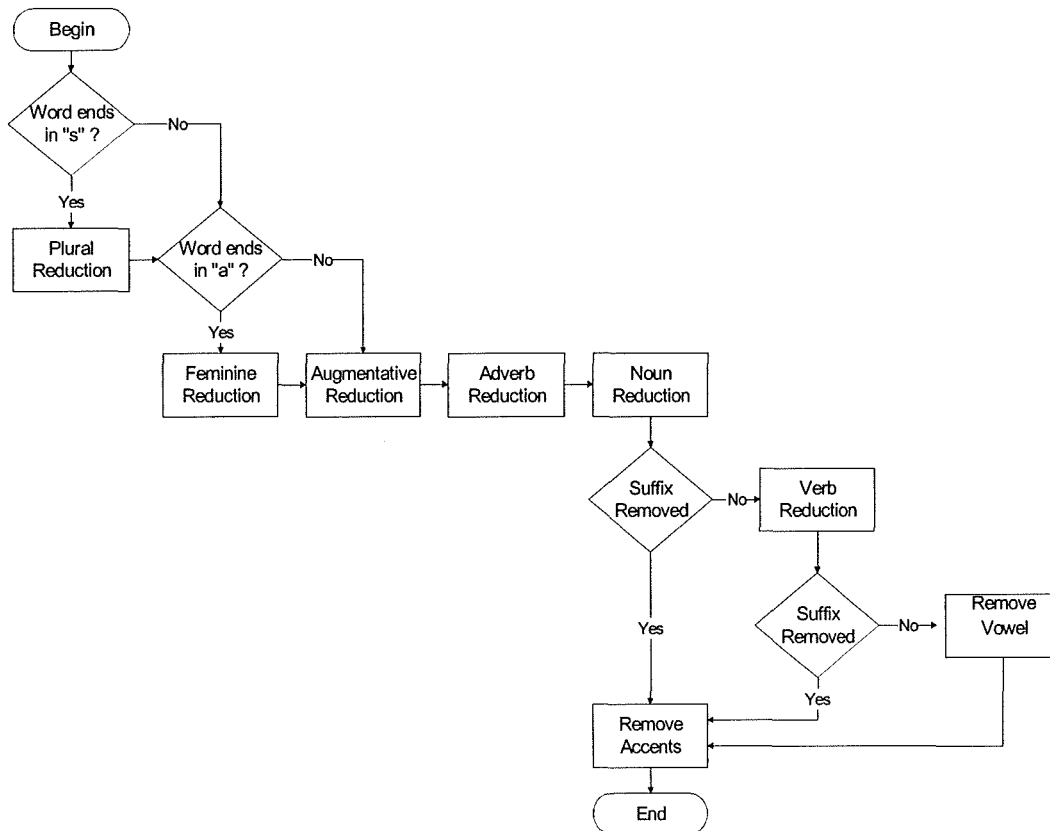


Figure 3.1 – Sequence of steps for the Portuguese Stemmer

The rules were created based on the most common Portuguese suffixes presented in [Cunha 1985; Macambira 1999]. The validity of such rules was checked using the evaluation methods presented in section 3.4. The Portuguese Stemmer contains 199 rules, please refer to the Appendix A for the complete list.

Each rule states:

- The suffix to be removed;
- The minimum length of the stem: this is to avoid removing a suffix when

the stem is too short. This measure varies for each suffix, and the values were set by observing lists of words ending in the given suffix. Although there is no linguistic support for this procedure, it reduces overstemming errors.

- A replacement suffix to be appended to the stem, if applicable;
- A list of exceptions: for nearly all rules defined, there were exceptions, so exception lists were added for each rule. Such lists were constructed with the aid of a vocabulary of 32,000 Portuguese words freely available from Snowball [Snowball]. Tests with the stemmer have shown that exceptions list reduce overstemming mistakes by 5%.

An example of a reduction rule is:

```
"inho", 3, "", {"caminho", "carinho", "cominho",  
"golfinho", "padrinho", "sobrinho", "vizinho"}
```

Where *"inho"* is a suffix that denotes diminutive, 3 is the minimum size for the stem, which prevents words like *"linho"* (linen) from being stemmed and the words between brackets are the exceptions for this rule; that is, they end in the suffix but they are not diminutives. All other words that end in *-inho* and that are longer than 6 characters will be stemmed. There is no replacement suffix in this rule.

Note that the stems do not have to be linguistically meaningful, since they are used to index a database of documents and are not presented to the user. They need however, to capture the word meaning without losing too much detail.

Step 1: Plural Reduction

With rare exceptions, the plural forms in Portuguese end in *-s*. However, not all words ending in *-s* denote plural, e.g. *lápiz* (pencil). This step consists basically in removing the final *"s"* of the words that are not listed as exceptions. Sometimes a

few extra modifications are needed e.g. words ending in *-ns* should have that suffix replaced by “*m*” like in *bons* → *bom*.

Step 2: Feminine Reduction

All nouns and adjectives in Portuguese have a gender. This step consists in transforming feminine forms to their corresponding masculine. Only words ending in *-a* are tested in this step but not all of them are converted, just the ones ending in the most common suffixes, e.g. *chinesa* → *chinês*.

Step 3: Adverb Reduction

This is the shortest step of all, as there is just one suffix that denotes adverbs – *mente*. Again not all words with that ending are adverbs so an exception list is needed.

Step 4: Augmentative/Diminutive Reduction

Portuguese nouns and adjectives present far more variant forms than their English counterparts. Words have augmentative, diminutive and superlative forms e.g. “small house” = *casinha*, where *-inha* is the suffix that indicates a diminutive. Those cases are treated by this step. According to [Cunha 1985] there are 38 of these suffixes, however some of them are obsolete. Therefore, in order to avoid overstemming, this algorithm uses only the most common ones that are still in frequent usage.

Step 5: Noun Suffix Reduction

This step tests words against 61 noun (and adjective) endings. If a suffix is removed here, steps 6 and 7 are not executed.

Step 6: Verb Suffix Reduction

Portuguese is a very rich language in terms of verbal forms. While the regular verbs in English have just 4 variations (e.g. talk, talks, talked, talking), the Portuguese regular verbs have over 50 different forms [Macambira 1999]. Each

one has its specific suffix. The verbs can vary according to tense, person, number and mode. The structure of the verbal forms can be represented as: root + thematic vowel⁴ + tense + person, e.g. *and* + *a* + *ra* + *m* (they walked). Verbal forms are reduced to their root as it would not be possible to reduce them to their infinitive without dictionary lookups.

Step 7: Vowel Removal

This task consists in removing the last vowel ("a", "e" or "o") of the words which have not been stemmed by steps 5 and 6, e.g. the word *menino* (boy) would not suffer any modifications by the previous steps, therefore this step will remove its final *-o*, so that it can be conflated with other variant forms such as *menina*, *meninice*, *meninão*, *menininho*, which will also be converted to the stem *menin*.

Step 8: Accents Removal

Removing accents is necessary because there are cases in which some variant forms of the word are accented and some are not, like in *psicólogo* (psychologist) and *psicologia* (psychology). After this step both forms would be conflated to *psicolog*. It is important that this step is done at this point and not right at the beginning of the algorithm because the presence of accents is significant for some rules e.g. *óis* → *ol* transforming *sóis* (suns) to *sol* (sun). If the rule was *ois* → *ol* instead, it would make mistakes like stemming *dois* (two) to *dol*.

3.3 Difficulties in Stemming Portuguese

Stemming Portuguese is much more troublesome than stemming English due to its more complex morphology. Below some particularly important difficulties are discussed:

⁴ There are 3 classes of verbs in Portuguese according to the ending of their infinitive form: "ar", "er", "ir". Thematic Vowel is the letter ("-a", "-e" and "-i") that groups verbs into categories.

- **Dealing with Exceptions:** As mentioned before one of the biggest difficulties in building this stemming algorithm was that for nearly every rule formulated there are exceptions, e.g. “*ão*” is a commonly used suffix to denote augmentative, however not all words ending in “*ão*” are in augmentative forms. Therefore, unlike Porter who opted for not having exceptions lists, the RSLP stemmer uses them. Not having exceptions lists would mean that the stemmer would make overstemming errors if the rule was kept or understemming errors if the rule was dropped.
- **Homographs:** There are several such cases, many involving conjugated verbs, e.g. *casais* which can mean “couples” or 2nd person plural of the verb “to marry”. The algorithm does not have information about word categories, so the different senses of those words are not distinguished. For this specific case, the stemmer assumes the first meaning and stems the word to its singular form *casal*, this is due to the 2nd person plural being nearly obsolete in modern Portuguese.
- **Irregular verbs:** The current version of the stemmer does not treat irregular verbs, but surprisingly they do not seem to seriously affect the results. The tests have shown that less than 1% of the mistakes occur because of this problem.
- **Changes to the morphological root:** There are cases in which the inflection process changes the root of the word. The cases in which the change obeys orthographic rules (e.g. *ns* → *m*) are being successfully treated. The other cases are still being incorrectly stemmed. At the moment, words like *emitir* (to emit) and *emissão* (emission), which are semantically related, are not being conflated, as the first is stemmed to *emit* and the later to *emis*.
- **Stemming proper names:** Proper names should not be stemmed, the problem is recognising them. A list of proper names does not present an ideal solution for two main reasons: there are infinite possibilities and some

proper names are shared by names of things, e.g. *Pereira* is a common Portuguese surname but it also means “pear tree”. Like the Porter stemmer, the present implementation of the algorithm is stemming proper names.

3.4 Evaluating The Stemmer

In order to assess the effectiveness of the stemmer a sequence of tests was carried out. Those tests used a vocabulary of 32,000 distinct word forms obtained from [Snowball]. The first test was to measure the time taken to stem that vocabulary. On a low end PC the time taken was 12 seconds. The algorithm was tested against the Portuguese version of the Porter stemmer applied to the same vocabulary. The main differences between the two stemmers are outlined below.

- Different morphology: the Portuguese version of the Porter Algorithm is a translation of the original English version, consequently it does not treat cases like augmentative and feminine forms since those do not occur in English. The RSLP algorithm was exclusively designed for Portuguese, so it deals with issues that are specific to Portuguese morphology.
- Accented characters: the Portuguese version of the Porter stemmer uses an intermediate representation for all accented characters (e.g. á is treated as a^a). The RSLP stemmer treats all accented characters as they are.
- Exceptions: the RSLP stemmer includes a list of exceptions for each rule as it was found they help avoiding overstemming errors.
- The steps composing the two algorithms are different.

The next sections describe the tests carried out using the two stemmers.

3.4.1 Vocabulary Reduction

One of the original purposes of suffix stripping was reducing the size of the vocabulary for indexing purposes. Porter reports a reduction of about a third of the vocabulary, using his stemmer over 10,000 different English words. The

Portuguese version of the Porter stemmer reduced the vocabulary by 44 %; this is because Portuguese has far more variant forms than English does. The RSLP stemmer reduces the vocabulary by 51%.

3.4.2 Comparison Against Expected Output

2800 words from the vocabulary were randomly selected, and for each one its correct stem was manually assigned. The calculated stems were then tested against the expected output and the errors were analysed to check if new rules or exceptions were needed. However, there is a stage in the development of a stemming algorithm when the addition of new rules with the target of avoiding errors on a specific case causes inaccuracies in other instances. At the end of the development process the RSLP stemmer was achieving 98% precision on the training vocabulary.

The stemmer was then evaluated using a set of words that had not been used in the training. Another 1000 words were selected, and again for each one its correct stem was assigned. The algorithm calculated the right stem 96% of the times outperforming the Portuguese version of the Porter stemmer which calculated the right stem 71% of the times, on the same test set.

3.4.3 Paice's Evaluation Method

One standard evaluation method of a stemming algorithm is to apply it to an IR test collection and calculate recall and precision to assess how the algorithm affects those measurements (see section 4.4.4 for an experiment). However, this method does not show the specific causes of errors therefore it does not help the designers to optimise their algorithms. Paice [Paice 1994] proposes an evaluation method in which stemming is assessed against predefined groups of semantically related words. He introduced new measurements: over and understemming indexes. This test requires a sample of different words partitioned into concept groups containing forms that are morphologically and semantically related to

one another. The perfect stemmer should conflate all words in a group to the same stem and that stem should not occur in any other group. Section 3.4.3.1 presents the method and section 3.4.3.2 provides a comparison between the RSLP stemmer and the Portuguese version of Porter algorithm.

3.4.3.1 Method Description

The method is applied as follows: For each concept group, two totals are computed:

- Desired merge total (DMT), which is the number of different possible word form pairs in the particular group, and is given by the formula:

$$DMT_g = 0.5n_g(n_g - 1) \quad (15)$$

where n_g is the number of words in that group.

- Desired Non-merge Total (DNT), which counts the possible word pairs formed by a member and a non-member word and is given by the formula:

$$DNT_g = 0.5n_g(W - n_g) \quad (16)$$

where W is the total number of words. The GDMT (global desired merge total) is obtained by summing the DMT for all groups and, similarly, by summing the DNT for all groups the GDNT (global desired non-merge total) is calculated. After applying the stemmer to the sample, some groups still contain two or more distinct stems, incurring understemming errors. The Unachieved Merge Total (UMT) counts the number of understemming errors for each group and is given by:

$$UMT_g = 0.5 \sum_{i=1..s} u_i(n_g - u_i) \quad (17)$$

where s is the number of distinct stems and u_i is the number of instances of each stem. The Global Unachieved Merge Total (GUMT) is calculated by summing the UMT for each group. The understemming index (UI) is given by: GUMT/GDMT.

After stemming, there can be cases where the same stem occurs in two or more different groups, which means there are overstemming errors. By partitioning the sample into groups that share the same stem the Wrongly Merged Total (WMT) is calculated. This measure counts the number of overstemming errors for each group. The formula is given by:

$$WMT_g = 0.5 \sum_{i=1..t} v_i (n_s - v_i) \quad (18)$$

where t is the number of original groups that share the same stem, n_s is the number of instances of that stem and v_i is the number of stems for each group t . By summing the WMT for each group the Global Wrongly-Merged Total (GWMT) is calculated. The overstemming index (OI) is given by: GWMT/GDNT.

Considering the following sample divided into 5 concept groups.

1. ajud (help): ajuda, ajudando, ajudinha, ajudei
2. duvid (doubt): duvido, dúvida, duvidamos, duvidem
3. chec (to check): checando, chequei, checamos, checou
4. beb (to drink): bebo, bebes, bebi, bebendo, bêbado, bebida
5. bebê (baby): bebê, bebezinho

Table 3.1 shows the calculation of GDMT and GDNT:

Group	DMT	DNT
1	6	32
2	6	32
3	6	32
4	15	42
5	1	18
Totals	34	156

Table 3.1 – Calculating GDMT and GDNT

After the stemming process:

1. ajud, ajud, ajud, ajud
2. duvid, duvid, duvid, duvid
3. chec, chequ, chec, chec
4. beb, beb, beb, beb, beb, beb
5. beb, beb

There are two kinds of mistake in this case: the words in group 3 were not conflated to the same stem (understemming), and words in groups 4 and 5 were stemmed to the same form (overstemming). Table 3.2 shows the results of GUMT:

Group	UMT
1	0
2	0
3	3
4	0
5	0
GUMT	3

Table 3.2 – Calculating GUMT

Now, the next step is to reorganise the sample into groups that share the same stem, in order to calculate GWMT (see Table 3.3):

1. ajud(1), ajud(1), ajud(1), ajud(1)
2. duvid(2), duvid(2), duvid(2), duvid(2)
3. chec(3), chec(3), chec(3)
4. chequ(3)
5. beb(4), beb(4), beb(4), beb(4), beb(4), beb(4), beb(5), beb(5)

Group	WMT
ajud	0
duv	0
chec	0
chequ	0
beb	12
GWMT	12

Table 3.3 – Calculating GWMT

The Understemming Index (UI) given by $GUMT/GDMT$ is 0.088 The Overstemming Index (OI) given by $GWMT/GDNT$ is 0.083. Paice's method does not offer a framework to assess the values of UI and OI, however it suggests that one stemmer will be better than another in terms of UI and worse in terms of OI. I disagree with that statement as one stemmer is better than the other if both OI and UI present lower values, that is, if it makes less overstemming and less understemming errors.

3.4.3.2 Comparison between the Two Stemmers

In order to compare the performance of the two Portuguese stemmers, another sample of 1000 words was selected and divided into 170 concept groups. The results obtained were:

- For the Portuguese version of the Porter stemmer:

UI: 0.215

OI: 2.11×10^{-4}

- For the RSLP stemmer:

UI: 0.034

OI: 9.85×10^{-5}

The results show that the RSLP stemmer makes less understemming errors and less overstemming errors than the Portuguese version of Porter's Algorithm, indicating that the stemmer proposed here is more effective.

Since its development, the RSLP stemmer has received a great number of requests. It is presently being used by researchers in over fifteen institutions as a basis for natural language processing applications. The rules described here have also been used as the foundation for the creation of a Galician stemmer [Brisaboa 2002].

3.5 Summary of the Chapter

This chapter presented the development of a stemming algorithm for Portuguese. The algorithm is simple yet highly effective and it is based on a set of steps composed by a collection of rules. Each rule specifies the suffix to be removed; the minimum length allowed for the stem; a replacement suffix, if necessary; and a list of exceptions.

The stemmer was evaluated using 3 different methods, including the one proposed by Paice [Paice 1994] and outperformed the Portuguese version of the Porter stemmer in all tests. The stemmer is applied in IR experiments described

in the next chapters. The algorithm has been implemented in C and is available at <http://www.cs.mdx.ac.uk/research/PhDArea/RSLP>.

4 PORTUGUESE-ENGLISH CROSS-LANGUAGE INFORMATION RETRIEVAL

This chapter presents initial experiments done using Portuguese queries to retrieve English documents. It begins by describing the resources used, then the experiments are reported, and finally some conclusions are presented. The Portuguese-English CLIR system described here is used as the basis for the experiments described in the next Chapter.

4.1 Why Portuguese?

Portuguese is the fifth biggest language in number of native speakers (see Figure 4.1). It is spoken in 4 continents: Europe (Portugal, Madeira, Azores), South America (Brazil), Africa (Angola, Mozambique, Cape Verde Islands, Guinea-Bissau) and Asia (Goa, Macau, East Timor). There are 176 million native speakers and another 15 million people use it as a second language.

In contrast, according to the Internet Software Consortium less than 1% of all web hosts are in Portuguese. In addition, only a small percentage of this population is competent in English, the vast majority (including graduates) are not able to present a query (in English) to a search engine like Google, Altavista, Lycos, etc. As a consequence, sources of information for the 8 million Portuguese speakers accessing the Internet are extremely limited, compared to the immense amount of information available to the English speaking population.

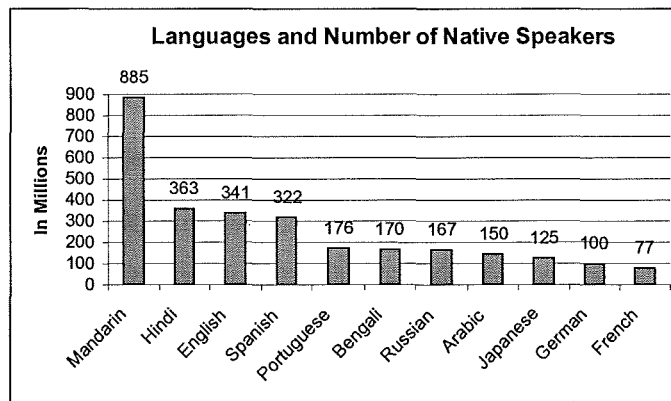


Figure 4.1– Languages in number of native speakers [Ethnologue]

In addition, Portuguese is very under-represented in terms of linguistic resources compared to other European languages such as French and Spanish. There is a lack of test collections, corpora and other freely available resources. Moreover, no CLIR research has been done using Portuguese. A valuable contribution that can be made to that community is to use Portuguese queries to retrieve English documents, and that is done in the experiments reported in this thesis. Such experiments can be carried out more easily by a Portuguese native speaker, and that is the case for the author of this work.

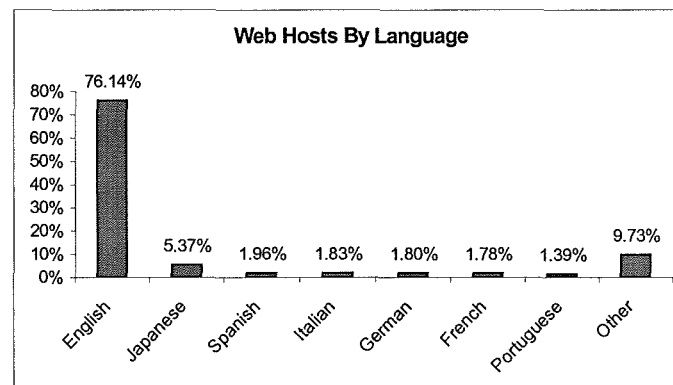


Figure 4.2 – Web hosts by language [Internet Software Consortium]

4.2 The CLIR System

LSI was chosen as the means to cross the language barrier (please refer to section 2.4 for an explanation of the method). This choice was mainly motivated by the good results achieved in thorough evaluation experiments [Yang 1997]. In addition, the technique does not demand expensive resources and can be easily

adapted to be used with new languages. Furthermore, performing relevance feedback on a LSI system is easier than in most IR approaches (see section 2.5.2). That feature is very important for the experiment described in Chapter 5.

4.3 Corpora

The ideal basis for a cross-language experiment using LSI is some parallel corpus. Unfortunately, there are no IR test collections containing parallel documents in English and Portuguese. Therefore, the alternative was to translate a sample of documents from English to Portuguese using SYSTRAN, thus creating a parallel collection. Four different test collections were used for the experiments:

- *Time*: articles from the Time Magazine world news section in 1963.
- *Cranfield 1400*: abstracts in aeronautics, originally used for tests at the Cranfield Institute of Technology in Bedford.
- *CISI*: abstracts in librarian science published between 1969 and 1977
- *LA Times*: news articles published during 1994.

The first three collections can be freely obtained from the University of Tennessee's website [*University of Tennessee*]. The LA Times collection was provided by CLEF [CLEF]. Details of these collections are given in Table 4.1.

	Time	Cranfield	CISI	LA Times
Number of documents	425	1400	1460	113005
Number of queries	83	225	112	50
Number of unique terms ⁵ – mono	7596	2629	3121	91190
Number of relevant docs	324	1838	3114	821

Table 4.1 – Details of the test collections

4.4 Experiments

For the small collections (Time, Cranfield and CISI) the first step consisted in translating half of the collection, i.e. all documents with odd identifiers were

⁵ Terms occurring in only one document were discarded.

translated. The queries were also translated. For the LA Times collection, only 20% of the documents were translated. This collection is composed of 365 files, one for each day of the year. One in every five files was picked for translation. CLEF [CLEF] provided a Portuguese version of the query topics; therefore translation was not need for the queries in this collection.

The second step was to remove the stopwords, according to the list provided by [Snowball]. The Porter stemmer was used to stem the English texts, and the RSLP stemmer described in chapter 3 was used on the Portuguese texts. Queries were also stemmed. Each document was then merged with its translation, creating a dual-language document.

Figure 4.3 shows a sample dual-language document used in the experiments. The translation is far from perfect, and many times the incorrect sense of a word was used, e.g. “branch” was translated to “filial” (shop branch), when the correct sense was “ramo” (tree branch). When SYSTRAN did not have a translation for a term, it remained in the original language. Nevertheless, no corrections were performed. Figure 4.4 shows the documents after stemming and stop-word removal.

<DOCNO> LA012394-0072 </DOCNO> TIME TO CARE FOR ROSES Give attention to roses at this time. Prune them before spraying to reduce the total area that needs to be sprayed at this time. Remember to drench the branches and trunk until the material runs off the branches .	<DOCNO> LA012394-0072 </DOCNO> A HORA DE IMPORTAR-SE COM ROSAS dá a atenção às rosas neste tempo. Pode-os antes de pulverizar para reduzir a área total que necessita ser pulverizada neste tempo. Recorde drench as filiais e o tronco até o material funciona for a das filiais .
---	---

Figure 4.3 – Sample dual-language document before stemming

<DOCNO> LA012394-0072 </DOCNO> time care rose give atten rose time prune sprai reduc total area sprai time rememb drench branch trunk materi run branch	<DOCNO> LA012394-0072 </DOCNO> hor import ros atenc ros temp pulver reduz are total pulveriz temp record drench filial tronc mater func filial.
--	---

Figure 4.4 – Sample dual-language document after stemming

The next stage was to derive the semantic space based on the dual-language documents, using the LSI software provided by Telcordia Technologies [Telcordia], who owns the patent to the technique. Then the remaining English-only documents were folded in, and finally the queries were executed. A hundred documents were retrieved in response to each query in all runs.

4.4.1 Weighting Scheme

The entries in the term by document matrix were, rather than the raw occurrences of terms in documents, the local weight (frequency of a term in a document) multiplied by the global weight of the term (number of occurrences of a term across the entire collection). Dumais [Dumais 1991], experimented with several weighting schemes in an LSI monolingual environment and concluded that the best weighting scheme is "log-entropy". This was the weighting scheme used for all experiments described here. It is given by the formula below. A term whose appearance tends to be equally likely among the documents is given a low weight and a term whose appearance is concentrated in a few documents is given a higher weight.

The elements of the matrix will be of the form: $L(i, j) * G(i)$

$$\text{Local Weighting: } L(i, j) = \log(tf_{ij} + 1) \quad (19)$$

$$\text{Global Weighting: } G(i) = 1 - \sum_{j=1}^N \frac{p_{ij} \log(p_{ij})}{\log N} \text{ where } p_{ij} = \frac{tf_{ij}}{gf_i} \quad (20)$$

where:

tf_{ij} = frequency of term i in document j

gf_i = total number of times term i occurs in the entire collection

N = number of documents in the collection

4.4.2 Number of dimensions

The first experiment aimed at finding a suitable number of dimensions for the next bilingual experiments. Because LSI does not offer any method for finding the optimal number of dimensions, several retrieval evaluations have been performed with a different number of dimensions. For the small collections 100 dimensions seems to work fine. However, for the LA Times collection an optimal number could not be established. The software used was not able to index the whole collection when more than 400 dimensions were used. For the experiments described in Chapter 5 only 20% of the collection was indexed, therefore it was possible to run them with more dimensions, however 700 was the maximum achieved.

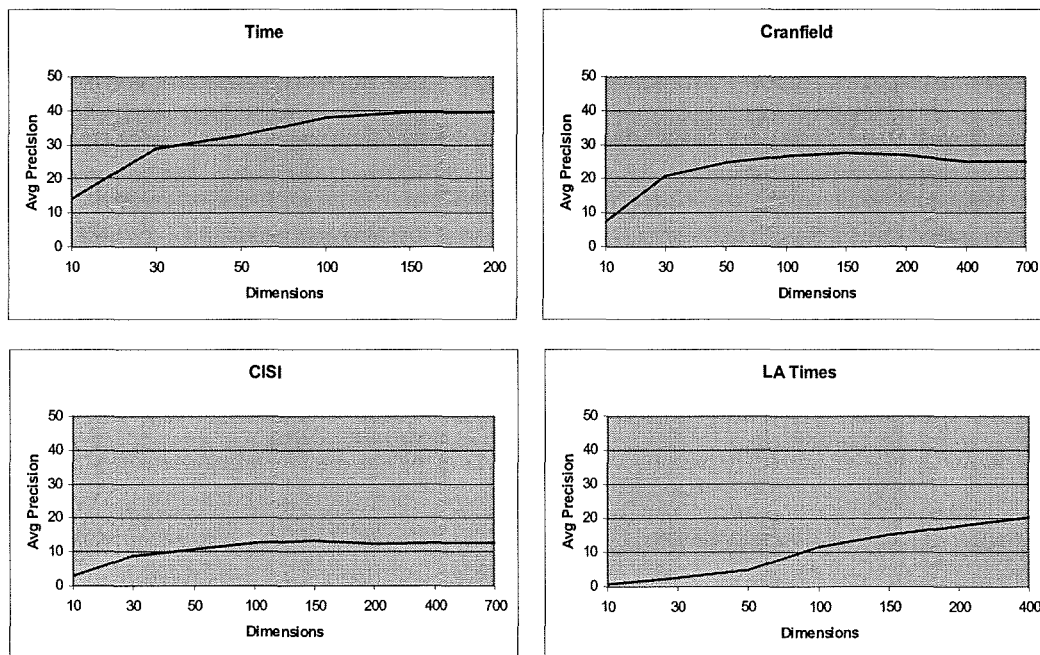


Figure 4.5 – Number of dimensions

4.4.3 Monolingual vs Bilingual Retrieval

The aim of this experiment was to assess the drop in performance suffered when another language is added to a LSI-IR system. Three runs are compared: (i) Monolingual – indexing all documents; (ii) Monolingual – Indexing half, folding in other half; and (iii) Bilingual. The results are summarised in Table 4.2.

Time			
	Mono – all	Mono – half	Bilingual
Number of Terms	2629	1979	8779
Average Precision	59.90%	38.13%	34.16%
Relevant Retrieved	311	311	299
Cranfield			
	Mono – all	Mono – half	Bilingual
Number of Terms	2629	1979	3429
Average Precision	29.53%	27.75%	26.60%
Relevant Retrieved	1441	1414	1382
CISI			
	Mono – all	Mono – half	Bilingual
Number of Terms	3121	2269	3875
Average Precision	15.31%	13.80%	12.56%
Relevant Retrieved	1240	1139	1097
LA Times			
	Mono – all	Mono – half	Bilingual
Number of Terms	91190	44932	69996
Average Precision	21.16%	20.34	17.09%
Relevant Retrieved	693	668	663

Table 4.2 – Summary of Results for Monolingual-Bilingual Comparison

Comparing the results of the mono-all run with the bilingual run, the loss in precision suffered ranges from 11% (Cranfield 29.53% → 26.60%) to 42% (Time Collection 59.90% → 34.16%). This represents the proportional loss, but the absolute decline ranged between 3% and 26%. However, this difference also includes the loss suffered due to the index in the bilingual system having fewer terms. That is a result of using only half of the collection for deriving the semantic space. Also, during the folding in process only new documents are added, not new terms. This fact is especially noticeable in the Time Collection

because it contains general domain documents comprising a larger vocabulary it suffers greater loss when fewer documents are used for indexing the database.

In order to assess the loss in performance incurred due to introduction of another language, the precision of the bilingual system should be compared with the precision of the monolingual system in which half the collection has been used for deriving the vector-space. Using this method, comparing the results from the mono-half and the bilingual runs, the proportional deficit ranges from 5% (Cranfield) to 19% (LA Times). The recall-precision curves shown in Figure 4.6 show that for three of the collections, all runs have very similar performances, which indicates that the use of MT to generate a parallel collection is feasible for the purpose of deriving the semantic space .

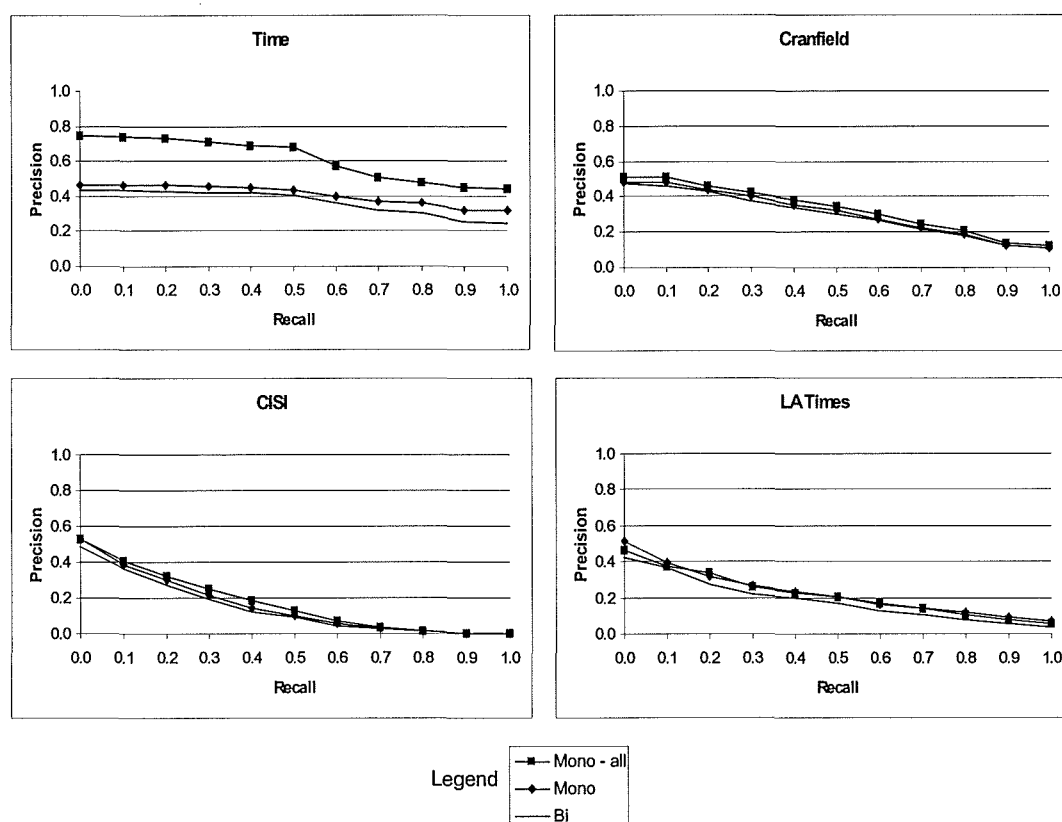


Figure 4.6 – Comparison between Monolingual and Bilingual IR

4.4.4 Effects of Stemming

The experiments reported in this subsection have mainly two aims: (i) assess the effects of stemming on different collections, and (ii) compare the performance of the RSLP stemmer proposed in Chapter 3 and the Portuguese version of the Porter stemmer.

Three runs of the experiment were performed: (i) Stemming the Portuguese texts using the stemmer proposed here; (ii) Stemming the Portuguese texts using the Portuguese version of the Porter Stemmer; and (iii) no stemming on any texts. A summary of the results is presented in Table 4.3 and recall-precision curves are presented in Figure 4.7. The tests with the small collections used 100 dimensions and the tests with the LA Times collections used 700 dimensions.

Time			
	Bilingual	Bilingual-Porter	No Stemming
Number of Terms	8779	8854	12049
Average Precision	34.16%	33.06%	31.61%
Relevant Retrieved	299	306	301
Cranfield			
	Bilingual	Bilingual-Porter	No Stemming
Number of Terms	3429	3624	5660
Average Precision	26.60%	25.97%	23.83%
Relevant Retrieved	1382	1368	1366
CISI			
	Bilingual	Bilingual-Porter	No Stemming
Number of Terms	3875	4116	6791
Average Precision	12.56%	12.29%	10.72%
Relevant Retrieved	1097	1069	950
LA Times			
	Bilingual	Bilingual-Porter	No Stemming
Number of Terms	69996	74163	96945
Average Precision	20.85%	16.46%	18.93%
Relevant Retrieved	694	658	671

Table 4.3 – Summary of Results for Stemming Experiment

Similar patterns can be observed in all tests. The RSLP stemmer achieved the highest results in all experiments: the average precision was higher and so was the number of relevant documents retrieved. However, the difference among the

runs was small. An average of 12% proportional improvement in relation to the Portuguese version of the Porter stemmer was observed in the small test collections, and a greater proportional improvement of 26% was observed in the LA Times collection.

No stemming performed the worst with the small collections and retrieved the least relevant documents in two out of the four collections. The results indicate that the use of stemming in a CLIR system through LSI is beneficial as it maximises the co-occurrences, providing better correspondence between cross-linguistic matches. The improvement achieved through stemming ranges between 8% and 17%.

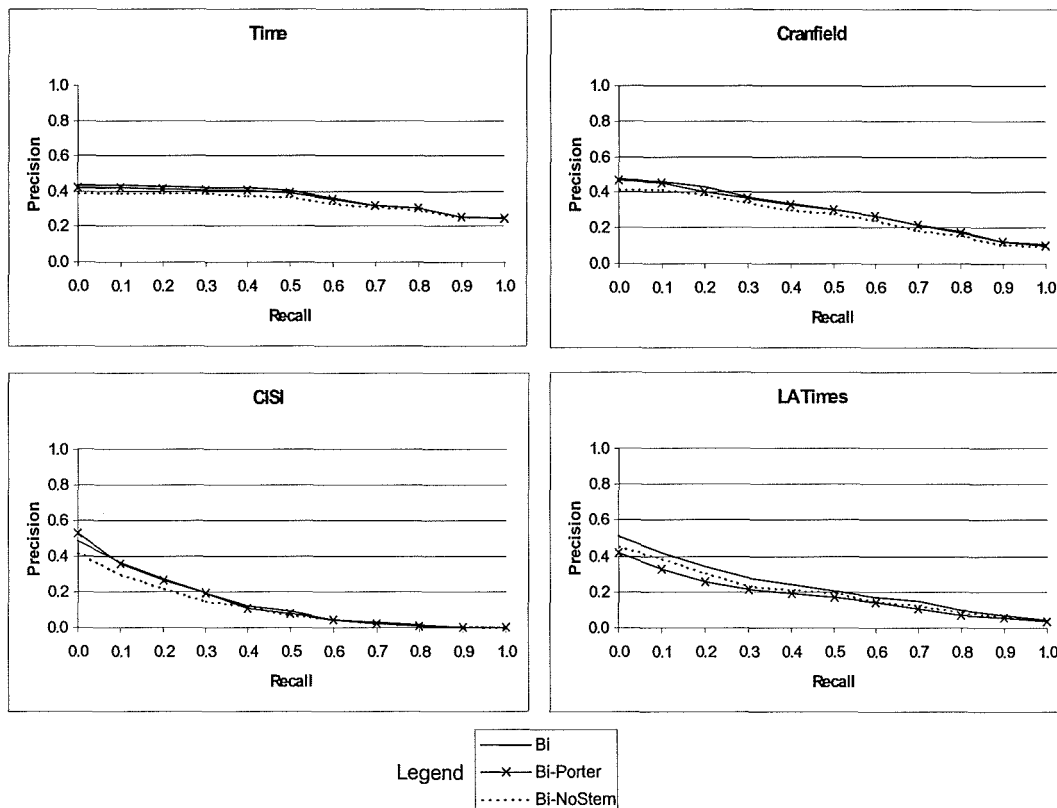


Figure 4.7 – Recall-Precision Curves for Stemming Experiment

4.4.5 Similarity between words

In an ideal situation the similarity between a term and its translation should be very high (close to 100%). In order to evaluate how accurately LSI represents the correspondence between terms, the similarities between 50 randomly chosen English words and their counterparts in Portuguese were calculated. Table 4.4 shows the results of this experiment for a sample of those words.

baby	bebê	99.67%	train	trem	12.30%
England	Inglaterra	99.24%	train	treinar	84.97%
eat	comer	94.19%	shade	sombra	31.88%
paper	papel	84.38%	shadow	sombra	31.88%
book	livro	77.11%	bank	banco	97.33%
Cyprus	Chipre	97.22%	bench	banco	32.94%
car	carro	91.28%	match	fósforo	96.72%
run	correr	51.03%	match	jogo	36.28%
find	encontrar	80.06%	game	jogo	91.30%

Table 4.4 – Term similarity across languages

The scores are generally quite high. However, when cross-linguistic polysemy is present, the similarity decreases significantly (see second column). This happens because term co-occurrences decrease in the presence of polysemy and results in a term and its translation being placed further apart in the concept space.

An important observation is the nearly perfect similarity between proper names. This is because even if proper names are translated, the mapping is almost always one to one (e.g. Germany is always translated to Alemanha). The same does not happen to nouns; polysemous terms can be translated to several different forms in another language (e.g. table can be translated to “*tabela*” or “*mesa*”, depending on whether it refers to the chart or to the piece of furniture). The poorest similarity occurred amongst verbs. In addition to suffering from polysemous cases, this category presents several irregular forms, which are not treated properly by either Portuguese or English stemmers. Stemming, however, is very important to verbs as it decreases the number of matches across different languages.

4.5 Discussion

The experiments in section 4.4.3 show that the bilingual execution of the system reaches 84 to 96% of the performance of the monolingual run. It seems plausible that some of that loss is due to the usage of MT to create a parallel corpus. It would have been interesting to assess the extent of such loss. However, that would require the manual translation of a few thousand documents, which is not feasible, given the resources available.

The experiment in section 4.4.5 shows that polysemous words pose a problem to LSI. A term is always represented by one vector only, even if it has several different senses. The term will be placed at the average of its different meanings. A potential problem would arise if none of the real meanings is near this average meaning. A possible solution to this problem would be to detect polysemous words, and place them in as many points in space as their number of meanings.

The results obtained with the bilingual version of the system for the LA Times Collection were average compared to the performance achieved by other systems in the CLEF campaign. During its four years of existence, results for bilingual experiments (involving queries in several languages) using this collection ranged from 1% to 47% average precision.

4.6 Summary of Chapter

This chapter reported on several CLIR–LSI experiments. The resources used were described. The experiments compare monolingual to bilingual retrieval, analyse the usage of stemming algorithms, and evaluate the ability of LSI to correctly map cross-linguistic synonyms. In summary, the main findings drawn from the experiments are:

- The cross-linguistic version of the system achieves 84% to 96% of the performance of a monolingual system.

- Stemming is beneficial, providing improvements ranging from 8% to 17%, and the RSLP stemmer proposed in chapter 3 is more effective than the Portuguese version of the Porter stemmer.
- The similarity of a term and its translation decreases in the presence of polysemy.

5 RELEVANCE FEEDBACK AND CROSS-LANGUAGE INFORMATION RETRIEVAL

Relevance Feedback (RF) acquires a new dimension in the context of Cross-Language Information Retrieval: the user's ability to recognise relevant documents written in a foreign language or translated into his language by some means. This chapter presents an experiment with two main aims. The first is to find out how well native Portuguese searchers can recognise relevant documents written in English, compared to documents that are hand translated and automatically translated to Portuguese. The second aim is to analyse the effect that misjudged documents have on the change in performance achieved through the RF process.

This chapter employs resources and applies concepts discussed in previous chapters. The stemmer presented in Chapter 3 is used by the Portuguese-English CLIR system discussed in Chapter 4, which serves as the basis for the RF experiment described here.

The next sections are organised as follows: section 5.1 presents the background for the experiments; section 5.2 describes the tests done to determine the RF method to be used; section 5.3 describes the experimental design; section 5.4 analyses user's ability to judge relevant documents on different systems; section 5.5 analyses the relationship between judgement errors and their effect on RF's

performance; section 5.6 presents the main findings; and section 5.7 summarises the chapter.

5.1 Background

RF has been widely applied to CLIR with positive results [Ballesteros 1997; Hiemstra 2000; McNamee 2002; Qu 2000; Yang 1997]. However, the vast majority of the experiments have used PRF or local context analysis (LCA); neither of these methods employ users to assess relevance.

Users' assessments of relevance are especially important for CLIR since the feedback process involves the subjects' ability to assess the relevance of documents written in foreign languages or automatically translated into the user's language.

This lack of experimental evidence has been addressed in part by the Interactive CLEF (iCLEF) [Oard 2001, 2003a, 2003b]. iCLEF provides a common framework for participant groups to evaluate several aspects related to the formulation of queries, translation of queries, and assessment of relevance.

The study developed by Karlgren and Hansen [Karlgren 2003] for iCLEF compared the performance of users assessing documents in their native language (Swedish) with their performance in assessing documents in a language they know well (English). As expected, they found that users take longer and make more mistakes when judging documents in a foreign language, even assuming a good knowledge of that language.

The use of MT in aiding relevance assessments was analysed by iCLEF in three studies:

- Wang and Oard [Wang 2001] compared the performance of full MT and term-for-term gloss translations obtained from bilingual term lists found on the web. Subjects had little or no knowledge of the language of the

documents. The results show that searchers were able to make relevance judgements with either approach. However, MT achieved slightly better results.

- Bathie and Sanderson [Bathie 2002] compared users' ability in judging native language documents and documents originally written in a foreign language and automatically translated into the user's language. The documents were articles from the LA Times in their original language and from Le Monde automatically translated into English. The study concluded that users were able to make judgements with the same accuracy for both types of documents.
- López-Ostenero et al. [López-Ostenero 2001] compared the performance of MT and a phrase translation based algorithm developed with the use of comparable corpora. Searchers had low or no proficiency on the language of the documents. The results show that precision was similar for both systems, but recall was better when using phrasal translations.

Though these experiments demonstrate that MT can facilitate relevance assessment in a CLIR environment, no study has yet examined the extent to which judging relevance may be better when using human translations of foreign language documents rather than MT. Further, the reviewed literature does not present any research on how the errors in judgement affect the change in performance achieved by the RF process. Those aspects are addressed by the experiment described in the next sections.

5.2 Choosing an RF Method

Most RF methods cannot be directly applied to CLIR as the words from the documents will not match words from the queries. The literature contains some cases in which RF methods have been adapted for CLIR. Yang et al. [Yang 1997] proposed a method for PRF on a bilingual collection which consists in doing an

initial retrieval on the collection that has the same language of the query; finding the translation mates for the top ranked documents; and then using those documents to create a query in the target language. Ballesteros and Croft [Ballesteros 1998] proposed a method for using PRF with dictionary methods. Three alternatives are tested: pre-translation query modification, post-translation query modification and a combination of both. Qu et al [Qu 2000], suggested a similar method for PRF in MT-based systems. All methods reported above mention performance improvements.

Whilst some methods must be adapted to work with CLIR, the same RF strategy used in a monolingual LSI system can be applied to a cross-lingual LSI system. That is possible because words are given a language-independent representation. Dumais [Dumais 1991] proposes three different methods for RF on a LSI system (those were presented here in section 2.5.2.4). In order to choose the best performing method to be used with the LA Times collection, four different RF methods were tested, two initially suggested by Dumais and two proposed here.

The four RF strategies attempted were:

- (RF1) Replacing the query with the highest ranked relevant document (proposed by Dumais).
- (RF2) Replacing the query with best match chosen by the user. This is a method tested here and tries to simulate the “more like this” feature which is present in some retrieval systems.
- (RF3) Retrieving documents closest to the centroid of relevant documents. This method was proposed by Dumais and has achieved the highest performance improvement with other (smaller) collections.
- (RF4) Expanding the original query with 20 terms picked from the best match. Harman [Harman 1992] recommends the use of 20 selected terms from the relevant documents. Based on that, this method selects the 20 most important terms from the document the user selected as the best match.

The RF procedure for all methods consisted in analysing for each query topic the top 20 ranked documents. In addition to selecting relevant items, a “best match” was also picked for each topic. The English version of CLEF topics for 2002 were used, and the documents returned were also in English. The set is composed of 50 topics, however 8 of them had no relevant documents in the collection, which leaves 42 queries to be analysed. For 10 topics none of the 20 top ranked documents analysed was considered relevant. This leaves 32 queries to be used for relevance feedback.

The results of this experiment are displayed in Figure 5.1. All RF methods presented an improvement in relation to the baseline. The best performing method was replacing the query with the centroid of the relevant documents (RF3), which confirms Dumais’ results. The methods RF1 and RF2 followed closely behind. The worst performing method was RF4 which yielded an improvement of 8%. It is worth pointing out that even with the best method, four topics had their performance worsened by the feedback process.

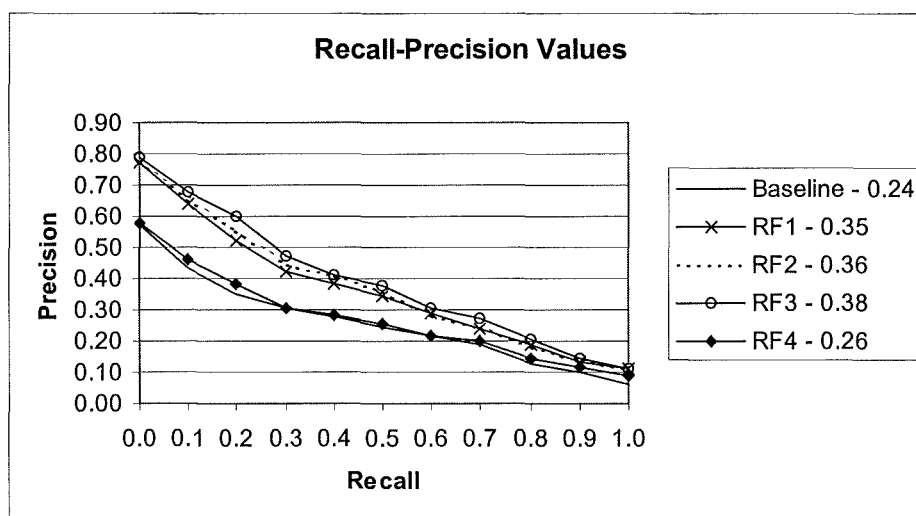


Figure 5.1 – Performance of RF methods for LSI

These tests not only formed the basis for the choice of the best RF method (RF3) but also revealed that analysing 20 documents per topic is too heavy a task for the subjects, especially when several topics are involved. This observation was

taken into account when designing the user experiment describe in the next section.

5.3 Experimental Design

The design of the experiment aims at answering two main questions:

- (i) how well can native Portuguese searchers recognise relevant documents written in English, compared to documents that are hand translated and automatically translated to Portuguese?
- (ii) what is the impact of misjudged documents on the performance improvement that can be achieved by RF?

The next sections describe the design of the experiment. Characteristics of the searcher, document collection, query topics and procedure are detailed.

5.3.1 Searcher

The aim was to obtain subjects that would be likely users of a CLIR system, in this case, Portuguese speakers who have basic or no knowledge of English, that are not able to express their queries in English and that are familiar with computer searching. The searchers were recruited among students and lecturers from UCPel (Universidade Católica de Pelotas⁶), in the south of Brazil. A total of 27 participants were obtained. The average age was 29.

Language skills are hard to measure accurately; what may be considered “intermediate” to one person, might be considered “advanced” by another. Ideally, the searchers would have taken a standard English language test such as TOEFL enabling a more exact categorisation of their knowledge. However, that was not possible. The approach taken was to ask the searchers to rate their ability in writing and reading in English (in two separate questions). There were 5 levels

⁶ <http://www.ucpel.tche.br>

of ability ranging from “none” to “proficient”. An example of question is presented in Figure 5.2. Most of the answers (13) fell into category 4. The remainder fell into categories 3 (8) and 5 (6).

How do you rate your ability of reading in English?

Proficient 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Unable

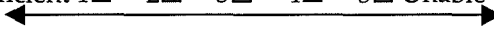


Figure 5.2 – English knowledge question

5.3.2 Document Collection and System

The collection used in the experiments consists of news articles from the Los Angeles Times published in 1994. Details of the collection are shown in Table 5.1.

	LA Times
Number of documents	113005
Number of queries	50
Number of unique terms ⁷ – bilingual	69996
Number of relevant documents	821

Table 5.1 – Details of LA Times

The CLIR system used is the bilingual system described in Chapter 4, in which 20% of the collection has been translated into Portuguese to simulate a parallel collection and serve as the basis for the generation of the semantic-space. The remaining 80% of English-only documents were then folded into that space. The number of dimensions used was 700.

5.3.3 Query Topics

Six query topics were extracted from CLEF 2002, which had a total of 50 queries. The Portuguese version of the topics was used. The criteria for topic selection was:

⁷ Terms occurring in only one document were discarded.

- Select topics that had more than 10 relevant documents. This criteria prevents the situation in which all relevant documents are presented to the user for feedback.
- Select topics that have relevant documents among top ten retrieved. Because the RF method used only positive feedback, this criteria was used to prevent the situation in which the user does not judge any document as being relevant.

Seventeen of the fifty topics satisfied the above conditions. Six of them were randomly selected. The English version of the selected topics is presented below:

Topic 1

```
<num> C092 </num>
<EN-title> U.N. sanctions against Iraq </EN-title>
<EN-desc> What measures has Iraq taken to effect the
lifting of the U.N. economic embargo and political
sanctions imposed after its invasion of Kuwait in 1990?
</EN-desc>
<EN-narr> Documents must include ways in which Iraq has
attempted to get the sanctions lifted. Mere
descriptions of the sanctions or rhetoric against the
sanctions are not relevant. Expressions of regret for
invading Kuwait by Iraqi officials are relevant. </EN-
narr>
```

Topic 2

```
<num> C094 </num>
<EN-title> Return of Solzhenitsyn </EN-title>
<EN-desc> Find documents which report about the return
of the Nobel prize winner for literature Solzhenitsyn to
Russia. </EN-desc>
<EN-narr> Relevant documents report the reasons and the
time of the return of Solzhenitsyn to Russia. They may
also talk about the reasons for his emigration to the
US. </EN-narr>
```

Topic 3

```
<num> C107 </num>
<EN-title> Genetic Engineering </EN-title>
<EN-desc> How does genetic engineering effect the human
food chain? </EN-desc>
<EN-narr> Articles must directly address the
introduction of genetic engineering, and its effects on
the human food chain. They will discuss both pros and
cons. Reports on tobacco bioengineering and human gene
engineering are not relevant. </EN-narr>
```

Topic 4

```
<num> C123 </num>
<EN-title> Marriage Jackson-Presley </EN-title>
<EN-desc> Find documents that report on the presumed
marriage of Michael Jackson with Lisa Marie Presley or
on their separation. </EN-desc>
<EN-narr> In May 1994, the famous pop star, Michael
Jackson, was reported to have married Lisa Marie
Presley, the daughter of the king of rock and roll.
Relevant documents must either contain some details
regarding the wedding, such as where or when it was
held, or must discuss the later separation of the
couple. </EN-narr>
```

Topic 5

```
<num> C130 </num>
<EN-title> Death of Nirvana leader </EN-title>
<EN-desc> How did the lead singer of the American rock
and grunge group, Nirvana, die? </EN-desc>
<EN-narr> Kurt Cobain, lead singer of Nirvana, the
famous popular music group, died in April 1994.
Documents that report the death of Cobain without
mentioning the cause are not relevant. </EN-narr>
```

Topic 6

```
<num> C140 </num>
<EN-title> Mobile phones </EN-title>
<EN-desc> Prospects for the use of cellular phones.
</EN-desc>
<EN-narr> Relevant documents report on the prospects for
the use of cellular phones and the development of the
mobile phone industry. </EN-narr>
```

5.3.4 Procedure

Searchers were presented with query topics (written in Portuguese) and a ranked list of 10 documents returned in response to an initial query. This ranked list was produced by presenting all terms from the “title” and “description” fields to the CLIR-LSI system, described in Chapter 4, as queries. The Portuguese version of the query topics was used. The documents whose vectors had the highest cosine with the query vector were ranked as best matches.

The participants were asked to classify each document in relation to the topic in one of three categories: “relevant”, “not relevant” or “not sure”. Similar to what is done by iCLEF [Oard 2001, 2003a, 2003b], the participants were given a definition of relevance. They were told to picture the situation in which they had to write a report on the query topic. They should consider relevant any document

that contains information on the topic. Documents in which only a part is related to the topic should be considered relevant. Additionally, each document should be judged independently of other documents, even if they contain the same information.

Each participant read 6 queries and 10 documents for each query, amounting to 60 relevance judgements per participant and 1620 in total. The users saw the full text of the documents, which was presented in one of the three formats presented below:

- the original English text (System 1)
- a machine translation produced using SYSTRAN (System 2)
- a human translation, produced by the author of this work⁸ (System 3)

The number of relevant documents per query varies. Similarly, the number of relevant documents ranked in the top ten and presented to the user varied. Table 5.2 shows the number of relevant documents per topic.

Topic	Number of relevant documents ranked in top 10	Number of relevant documents in the collection
1	6	27
2	8	12
3	5	32
4	2	19
5	7	63
6	5	70

Table 5.2 – Number of relevant documents per topic

The order of the queries was varied systematically in a Latin square design, which controlled for the learning effect and tiredness of the searchers. The order in which the different systems were presented has also been varied. Table 5.3 shows a 9-subject matrix. As there were 27 participants, the same matrix was used three times. Participant 1 saw the documents for topics 1 and 2 in the

⁸ The author is a Portuguese native speaker and fluent in English.

original language (English), then the documents for topics 3 and 4 automatically translated into Portuguese and finally, documents for topics 5 and 6 manually translated to Portuguese. Participants 1, 2 and 3 had the same topic-system combination, however the order in which the query topics were presented was different for each subject. The average time taken to judge all sixty documents was one hour.

Participant	First Batch			Second Batch			Third Batch		
1	S1	1	2	S2	3	4	S3	5	6
2	S2	3	4	S3	5	6	S1	1	2
3	S3	5	6	S1	1	2	S2	3	4
4	S1	3	4	S2	5	6	S3	1	2
5	S2	5	6	S3	1	2	S1	3	4
6	S3	1	2	S1	3	4	S2	5	6
7	S1	5	6	S2	1	2	S3	3	4
8	S2	1	2	S3	3	4	S1	5	6
9	S3	3	4	S1	5	6	S2	1	2

Table 5.3 – Subject matrix showing topic-system combination and the presentation order

Besides providing relevance judgements, the users were also asked some questions. Before the experiment they were asked about their language skills and experience in computer searching. During the experiment, after judging the documents for each query topic, they were asked about their confidence in the judgements made, their prior knowledge of the topic, and how difficult they found making the judgements. At the end of the experiment they were asked if they preferred to view the documents in their original language or translated into Portuguese. Finally, they were asked to give any comments they would like to make. A sample questionnaire is presented in Appendix B, along with the documents judged for relevance. The data collected with this experiment, along with the statistical tests performed is presented in Appendix E.

After gathering the judgements for all 27 searchers, the queries were reformulated, resubmitted and re-evaluated for recall and precision. The RF method used was RF3, which was discussed in section 5.2.

5.4 Users Ability in Making Relevance Judgements

As reported in the previous section, a three-point relevance scale was used. However, to be compatible with CLEF assessments and the evaluation software, all “not sure” were forced to “irrelevant”. The analysis concentrated mainly on the following aspects:

1. The number of mistakes committed by the searcher.
2. The level of agreement between the CLEF judgements and the judgements of each user.
3. The confidence in the judgements made, the difficulty of the task, and prior knowledge of the topics.

The data collected for most variables is not perfectly normally distributed. The statistical test chosen to compare the results for different groups was ANOVA as it is robust in dealing with data that depart from the normality assumption. For all tests reported by in the next subsections, α^9 was set to 0.05.

5.4.1 Number of Mistakes

The relevance judgements provided by CLEF were considered as “correct answers”. Each judgement collected from the participants was compared against them. Two types of mistakes were analysed: (i) false alarm, if the searcher judged an irrelevant document as relevant; and (ii) relevant missed, if the user judged a relevant document as irrelevant.

A total of 1620 judgements were made, 540 for each system. Because each query topic had a variable number of relevant documents, the figures for relevant

⁹ α measures the probability of a Type 1 error, which occurs if the Null Hypothesis is rejected when, in fact, it should be accepted. The Null Hypothesis for ANOVA tests states that there is no significant difference amongst means taken from different samples.

missed and false alarm had to be properly weighted to allow for fair comparisons between topics.

Table 5.4 shows the number of missed relevant, the number of false alarms, and the number of correct judgements. It also displays how the judgements spread across the 3 possible categories: relevant, not relevant and not sure.

	Hand Translated	Machine Translated	Original
Missed Relevant	142 (26%)	141 (26%)	234 (43%)
False Alarm	94 (17%)	80 (15%)	43 (8%)
Correct	304(56%)	319(59%)	263(48%)
Relevant	249 (46%)	236 (43%)	106 (19%)
Not Relevant	243(45%)	241(45%)	96(18%)
Not Sure	48 (9%)	63 (12%)	338 (63%)

Table 5.4 – Summary of Judgements

The number of relevant missed has been virtually the same for the machine translated texts and the hand translated texts. The number of relevant missed for the original texts was much higher (43%). That happened because most judgements (63%) for this system fell into the unsure category. The number of false alarm has been very small in the original texts for the same reason, and has been the largest for the hand translated texts because people made more positive judgements in that system. An ANOVA test on missed relevant and false alarm has shown no significant difference between judgements made using hand and machine translated texts (p -values 0.99 and 0.36 respectively – see page 240-241).

5.4.2 Overlap

Overlap has been defined by Lesk and Salton [Lesk 1968] as the intersection of the relevant documents divided by the union of the relevant document sets. In the context of this experiment, overlap measures how accurate the participant's judgements were, as it tells how similar each participant's judgements were compared to the relevance judgements provided by CLEF.

Only the documents presented to the users were considered when calculating the overlap. For example, considering the following situation presented in Table 5.5. The intersection of the relevant documents is 2 as the user's judgements agree with the CLEF judgements in two cases, for documents 1 and 5. The union of the relevant documents is 9 as there are 5 documents that were considered relevant by CLEF judges and 4 documents by the user. Therefore the overlap is 0.22.

Document	CLEF Judgements	User's Judgements
1	relevant	relevant
2	not relevant	relevant
3	relevant	not relevant
4	relevant	not relevant
5	relevant	relevant
6	not relevant	not relevant
7	not relevant	not relevant
8	not relevant	not relevant
9	not relevant	relevant
10	relevant	not relevant

Table 5.5 – Overlap Example

The average overlap between the CLEF assessors and the participants of the experiment is shown in Table 5.6. These results confirm the ones from the previous section, as the figures for the hand translated and machine translated texts are very similar, and much better than the scores achieved when assessing the original texts. An ANOVA test has shown no significant difference (p -value = 0.85 – see page 242) between the overlap of judgements made on hand and machine translation texts.

	Overlap
Hand Translated	0.40
Machine Translated	0.41
Original	0.16

Table 5.6 – Overlap by System

5.4.3 Difficulty, Confidence and Knowledge

The participants were asked to rate each topic in terms of difficulty of the task and the confidence they had in the judgements made. They were also asked how

familiar they were with the subject of the query topics. For all three questions there were five levels. An English version of the questions asked is shown in Figure 5.3.

How difficult did you consider the task of judging the document for this topic?
 Very Easy 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Very Difficult

How confident are you about the judgements you made ?
 Very Confident 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Very Unsure

How familiar are you with the subject of this topic?
 Expert 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ No Knowledge

Figure 5.3 – Questions on difficulty, confidence and familiarity

Participants found judging English documents considerably more difficult than judging MT or hand translated documents. They also had less confidence in the judgements made. Confidence and difficulty were similar for MT and hand translated documents. Figure 5.4 and Figure 5.5 show histograms of the frequency that each answer was given.

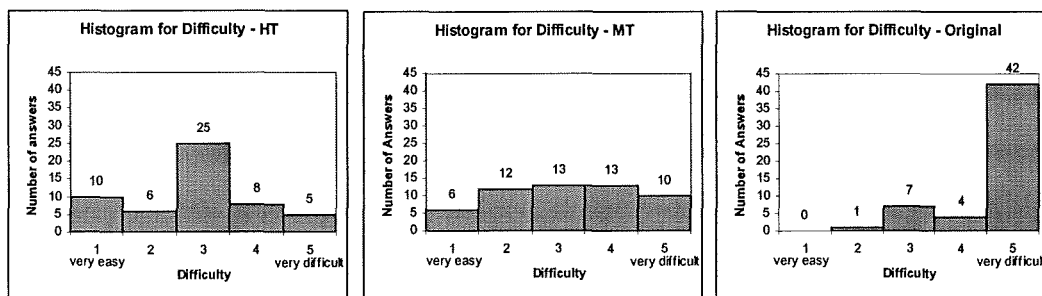


Figure 5.4 – Histogram for Difficulty of the Task (per system)

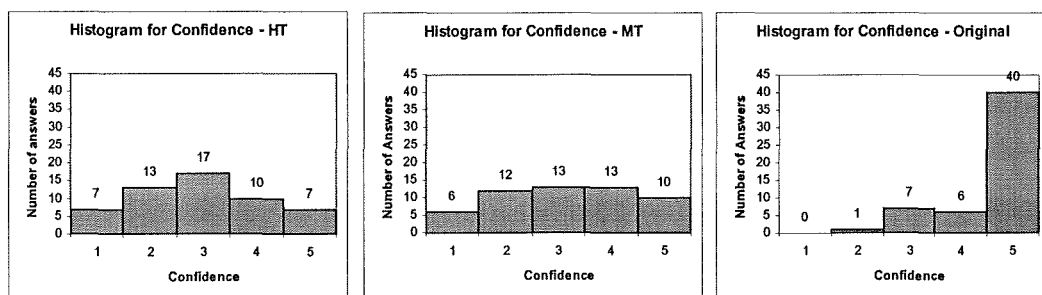


Figure 5.5 – Histogram for Confidence in Judgements (per system)

Those three measures have also been evaluated separately for each topic. Difficulty and confidence had very similar distributions along the five categories. No significant difference was found among the level of difficulty of the topics. Likewise, no significant difference was found among the degree of confidence in the judgements made for each topic. An ANOVA test confirms that, the calculated p -values are 0.58 when comparing the difficulty of the 6 topics, and 0.79 when comparing the level of confidence in the judgements (see page 246).

Prior knowledge of the topic had a greater variation. Topic 2, about Solzhenitsyn, was the least known. An ANOVA test has shown that these differences are not significant (p -value=0.11). Figure 5.6 shows histograms for the frequency of each answer for the question on the knowledge of the topic.

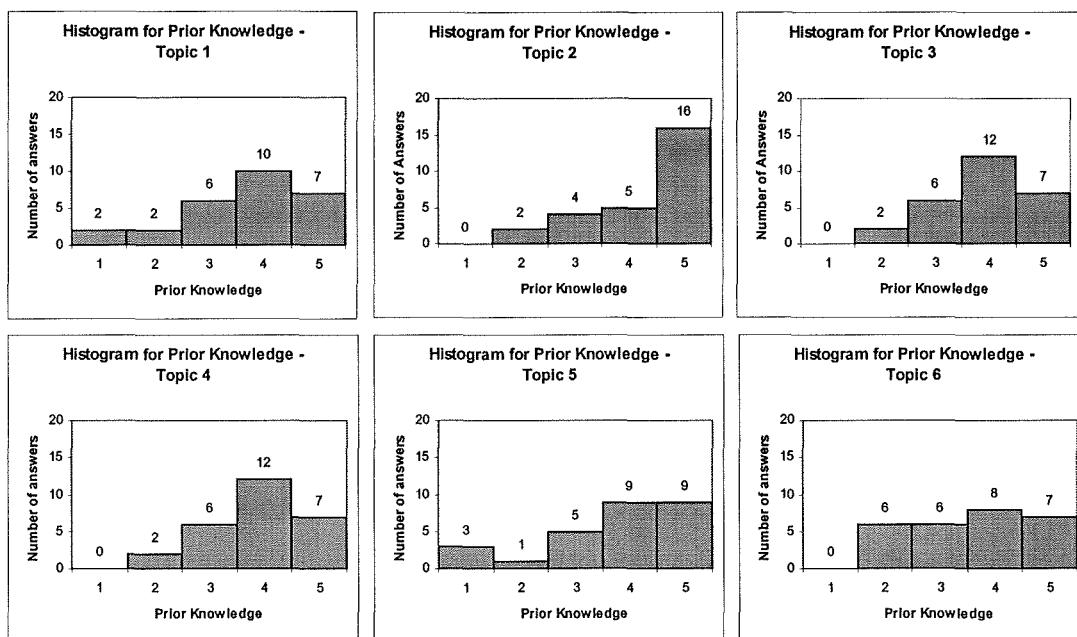


Figure 5.6 – Histograms for Prior Knowledge of the Topics

A moderate correlation of 0.34 was found between the confidence on the judgements and the overlap. And a weaker correlation of 0.27 was found between the subjects' knowledge of the subject of the query topic and the overlap. Both coefficients are significant at a 99% confidence level (see page 249). This relationship however, is not very strong. Squaring the correlation coefficient

(*r*), it can be concluded that only 11% of the overlap can be explained by the confidence in the judgements and 7% by the knowledge of the subject.

5.4.4 Testing Other User Groups

In order to verify the results and provide means for comparison, the experiment was repeated with two other groups of participants: one group composed of native Portuguese speakers with a very good command of English and a group of native English speakers. Those tests are described in the next two subsections

5.4.4.1 Testing Bilingual Users

The experiment was repeated with a smaller sample of 6 participants, all native Portuguese speakers with very good English skills. Four participants rated their ability in reading and writing in English on category 1 (proficient) and the remaining two rated their ability on category number two. A total of 36 queries were analysed, 12 for each system.

The results are summarised in Table 5.7. Bilingual participants were able to make judgements with the same accuracy using hand translated, machine translated and original texts. An ANOVA test using the data from false alarm, missed relevant and the overlap showed no significant difference among the three means. Complete data is presented on page 237.

	Hand Translated	Machine Translated	Original
Missed Relevant	21 (18%)	25 (20%)	30 (25%)
False Alarm	11 (9%)	19 (16%)	21 (18%)
Correct	88 (73%)	76 (64%)	69 (57%)
Relevant	66 (55%)	67 (56%)	67 (56%)
Not Relevant	51 (43%)	34 (28%)	45 (38%)
Not Sure	3 (2%)	19 (16%)	8 (6%)
Overlap	0.55	0.49	0.43

Table 5.7 – Results for Bilingual Users

5.4.4.2 Testing Native Speakers

The aim of testing native English speakers was to establish the expected degree of agreement between a participant and CLEF judges, when the participant fully understood the language of the documents. The 6 participants recruited saw only the English documents, and each judged 2 queries (20 documents). In total, data for 12 queries were analysed. Table 5.8 shows the results for this group. Complete data is presented on page 239.

Missed Relevant	36 (30%)
False Alarm	3 (3%)
Correct	81 (67%)
Relevant	33 (28%)
Not Relevant	78 (65%)
Not Sure	9 (7%)
Overlap	0.46

Table 5.8 – Results Native Speakers

The results obtained here are comparable to the ones obtained by Voorhees [Voorhees 1998] in an experiment using three groups of relevance assessors, all native English speakers from a similar background judging English documents. She found that the overlap between pairs ranged from 0.42 to 0.49.

5.4.5 Comparison Among Groups

Table 5.9 shows a summary of the results for accuracy in judgement for all three groups of users: Portuguese speakers with poor English skills, bilingual subjects, and native English speakers. The performance of the Portuguese participants from both groups judging hand and automatically translated documents is equivalent to the performance of native English speakers judging English documents. An ANOVA test has shown no significant difference amongst those groups in terms of overlap (see page 245). In fact, the only group that shows a difference in relation to the others was monolingual Portuguese users judging English documents.

	Portuguese speakers with poor English skills			Bilingual Subjects			Native English Speakers
	HT	MT	Original	HT	MT	Original	Original
Missed Relevant	26%	26%	43%	18%	20%	25%	30%
False Alarm	17%	15%	8%	9%	16%	18%	3%
Correct	56%	59%	48%	73%	64%	57%	67%
Relevant	46%	43%	19%	55%	56%	56%	28%
Not Relevant	45%	45%	18%	43%	28%	38%	65%
Not Sure	9%	12%	63%	2%	16%	6%	7%
Overlap	0.40	0.41	0.16	0.55	0.49	0.43	0.46

Table 5.9 – Comparison among groups

An ANOVA test of missed relevant confirmed the result obtained when testing overlap (see page 243). In contrast, a significant difference was found when testing false alarm between the English speakers and the Portuguese speakers with poor English skills judging hand-translated documents (see page 244). The English participants made fewer mistakes of that type (3%) compared to the other group judging hand translated documents. This indicates that native English speakers were stricter in judging a document as relevant. That fact has apparently influenced the slightly higher number of missed relevant documents they had in comparison to the other groups. However, the difference was not significant for that variable.

5.5 Errors in Judgement and Change in Performance

The precision of the initial query was compared with the precision attained after the relevance feedback process for each user. This measure has been taken using the "residual collection" method (please refer to section 2.5.3 for a description of the method).

Table 5.10 shows average precision for the baseline run and different feedback runs using different sets of relevance judgements. The feedback runs compared are:

- CLEF – using the official judgements provided by CLEF.

- Best – using the judgements provided by the participant who achieved the biggest overall improvement.
- Worst – using the judgements provided by the participant who got the largest overall decline.
- Optimal – selecting the set of judgements that yielded the best result for each topic.
- Average – combining the results from all participants for each topic.

Topic	Baseline	Official	Best	Worst	Optimal	Average
1	0.4633	0.7176	0.7287	0.4633	0.7393	0.6517
2	0.2665	0.0704	0.0722	0.0273	0.2665	0.1287
3	0.1064	0.0557	0.4773	0.0039	0.6012	0.1256
4	0.0599	0.1813	0.1813	0.0570	0.2617	0.1366
5	0.1661	0.2171	0.2631	0.1748	0.2779	0.2021
6	0.4694	0.4831	0.6457	0.4652	0.6928	0.4792
Average	0.2553	0.2875	0.3947	0.1986	0.4732	0.2873
Change	-	+12.64%	+54.62%	-22.20%	+85.38%	+12.53%

Table 5.10 – Average precision figures for initial and feedback runs

The change in performance varied greatly from one user to another, ranging from a deterioration of 22% to an improvement of 54% (using the residual collection). Figure 5.7 shows the histogram for change in precision; it clearly illustrates that most participants achieved a small improvement. Averaging the results for all users, there was an improvement of 12.53% on mean average precision. Using the relevance assessments provided by 6 participants the feedback runs performed worse than the baseline. An important observation is that the official judgements, which are the gold standard, did not produce the biggest overall improvement. In fact, the official run was outperformed by the runs of 10 participants of the main experiment, which implies that judgement errors may sometimes help the RF process. That aspect is discussed in greater detail in section 5.6.3. It is possible that if using a different IR system, the result achieved with the different feedback runs would be different. One possible reason for the poor performance of CLEF judgements in comparison to user judgements is that some algorithms achieve better performance in the presence of noise.

Figure 5.8 shows the recall-precision curves for the baseline and feedback runs presented on Table 5.10.

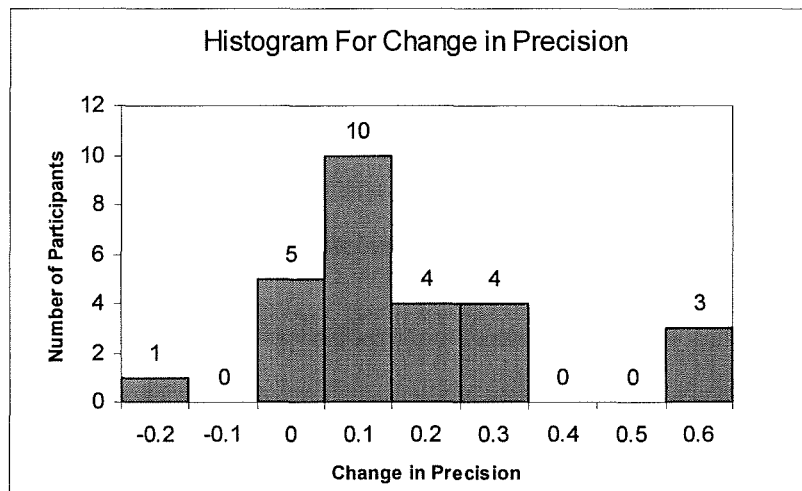


Figure 5.7 – Histogram for change in precision (by user)

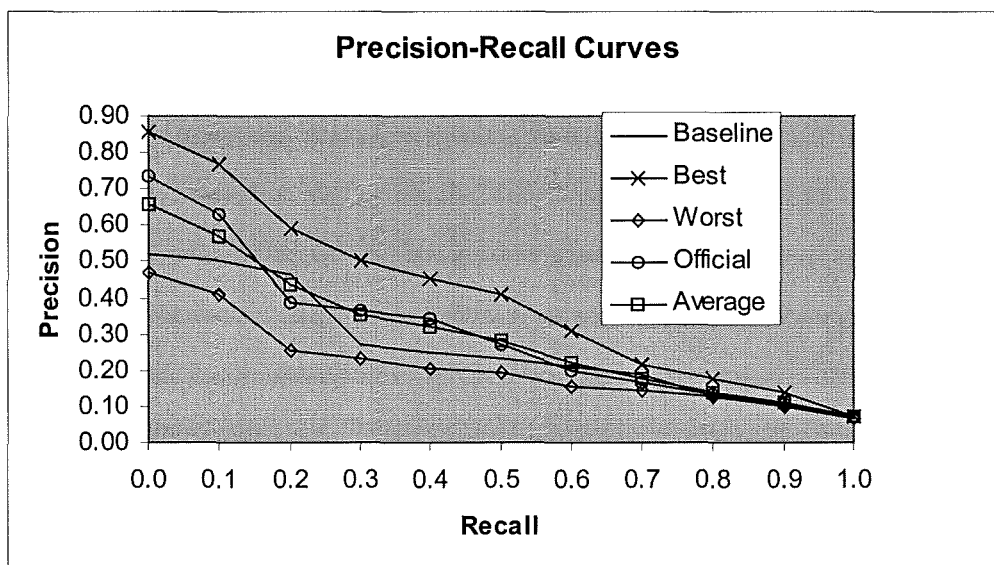


Figure 5.8 – Recall–Precision curves using the residual collection.

A ranking of participants in terms of change in performance and overlap is presented in Table 5.11. There is a clear indication that the improvement provided by relevance feedback is not determined by the accuracy in judging relevance.

Sorted in Decreasing order of Change in Precision			Sorted in Decreasing order of Overlap		
Participant	Overlap	Change	Participant	Overlap	Change
5a	0.5128	0.5463	9c	0.6053	0.2486
3b	0.3824	0.5319	5a	0.5128	0.5463
5c	0.3415	0.5195	7a	0.4884	0.1312
2a	0.3659	0.2752	8b	0.4762	0.2285
9c	0.6053	0.2486	2b	0.4545	0.1247
8b	0.4762	0.2285	4b	0.4545	0.0513
1a	0.1944	0.2281	6c	0.4186	0.0849
7c	0.3333	0.1927	7b	0.3864	0.0524
4c	0.3000	0.1312	3b	0.3824	0.5319
7a	0.4884	0.1312	2a	0.3659	0.2752
2b	0.4545	0.1247	8a	0.3529	-0.0193
1b	0.2500	0.0930	9b	0.3469	0.0649
6c	0.4186	0.0849	6b	0.3469	0.0249
6a	0.2500	0.0799	5c	0.3415	0.5195
5b	0.2973	0.0693	7c	0.3333	0.1927
9b	0.3469	0.0649	9a	0.3333	0.0546
9a	0.3333	0.0546	4c	0.3000	0.1312
7b	0.3864	0.0524	5b	0.2973	0.0693
4b	0.4545	0.0513	8c	0.2571	0.0468
8c	0.2571	0.0468	1b	0.2500	0.0930
6b	0.3469	0.0249	6a	0.2500	0.0799
3a	0.1667	-0.0003	4a	0.2500	-0.0422
8a	0.3529	-0.0193	3c	0.2105	-0.0712
2c	0.1190	-0.0368	1c	0.2000	-0.2221
4a	0.2500	-0.0422	1a	0.1944	0.2281
3c	0.2105	-0.0712	3a	0.1667	-0.0003
1c	0.2000	-0.2221	2c	0.1190	-0.0368

Table 5.11 – Participants ranked by change and overlap

In order to establish what affects the performance of RF, it is necessary to analyse the effect that each variable had in the change in performance achieved. This can be done by calculating correlation coefficients between the change in performance and the other variables analysed. The correlation coefficients obtained when analysing the results for all 162 queries are shown in Table 5.12 and presented in more detail in Appendix E page 249.

	Change in Precision
Missed	-0.24
False Alarm	-0.27
M+F	-0.51
Overlap	0.33

Table 5.12 – Correlation between change in precision and accuracy in judgement

The overall results show a moderate negative correlation between the change in performance and missed relevant. A similar moderate negative correlation is found between the change in performance and false alarm. That is somewhat surprising as it seems logical that selecting an irrelevant document as relevant is a more serious mistake than failing to recognise a relevant document, however this trend seems to be very small. Similarly, there is a moderate positive correlation between the change in performance and the overlap. In summary, squaring the correlation coefficients, it can be said that 7% of the change in performance can be explained by the false alarms, 6% by the missed relevant and 11% by the overlap or accuracy in judgement.

5.5.1 Evaluation by System

With the purpose of understanding if the different systems had an effect on the performance change, the next step is to compare the results for different systems. Table 5.13 shows correlation coefficients between the improvement in precision with the misjudged documents (relevant missed and false alarm), and with the overlap for each system. For correlations with other variables please refer to Appendix E pages 250-252.

	Change in Precision Hand Translation	Change in Precision Machine Translation	Change in Precision Original
Missed	-0.12	-0.25	-0.38
False Alarm	-0.37	-0.27	-0.22
M+F	-0.53	-0.47	-0.54
Overlap	0.25	0.31	0.49

Table 5.13 – Correlation between change in precision and accuracy in judgement (by system).

The sum of mistakes has very similar coefficients for all three systems. Differences are found for false alarm, as it is higher for hand translation. That can be attributed to this type of mistake being more frequent for that system than for the other two. Likewise, missed relevant had a higher coefficient in the original texts because that system suffered more from that type of mistake. However, the

differences among systems are not large enough to justify assigning the effect of the changes in performance to system variations.

5.5.2 Evaluation by Topic

The next step then is to evaluate the correlation coefficients topic by topic. Appendix E shows the data from the user experiment, Table 5.14 extracts some important correlation coefficients between change in precision and the mistakes for each of the six query topics. For correlations with other variables please refer to Appendix E pages 253-258.

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6
Missed	-0.83	0.55	-0.05	-0.59	-0.56	-0.21
False Alarm	0.67	-0.61	-0.48	-0.26	0.02	0.10
M+F	-0.59	0.41	-0.47	-0.51	-0.61	-0.16
Overlap	0.86	-0.52	0.24	0.70	0.61	0.25

Table 5.14 – Correlation between change in precision and accuracy in judgement (by topic).

The data shows big differences from one topic to another. A summary of the findings for each topic is presented below:

- Topic 1: All feedback runs achieved some improvement compared to the baseline. The improvement in performance was strongly correlated with overlap. Amazingly, false alarms had a positive correlation with improvement in performance.
- Topic 2: No feedback run managed to achieve any improvement in relation to the baseline. Performance decreased as much as 88% (see Appendix E page 237). There were two surprises, overlap is negatively correlated ($r = -0.52$) and relevant missed is positively correlated with the improvement in performance.
- Topic 3: There was a big gap between the best and worst feedback runs (+465% and -99% - see Appendix E page 232). There was nearly no correlation between improvement in performance and missed relevant, and a moderate positive correlation with overlap.

- Topic 4: Most runs achieve great performance improvements. Overlap was strongly positively correlated with change in performance.
- Topic 5: Nearly all feedback runs achieved improvements. There was no correlation between false alarm and the improvement in performance.
- Topic 6: There was an even split between gain and deterioration in performance. Missed relevant, false alarm and overlap presented weak correlations with change in performance.

5.6 Findings

This section presents the findings drawn from the experiments. The questions proposed before the experiment (see section 5.3) are answered, and other pertinent points are explored.

5.6.1 RF improves the performance of a CLIR–LSI system

Despite all differences in judgements, RF has proven to be beneficial for most cases. The mean average precision for the 6 topics analysed was improved 77% of the time (21 out of 27 participants). The average improvement was of 12% (see section 5.5).

Analysing the results topic by topic the RF process achieved an improvement in 5 of the 6 topics when averaging results from all participants. Only query 2 had its performance decreased. Query 4 was the one that most benefited from RF. A moderate negative correlation ($r=-0.40$ – see page 249) was found between initial average precision of the queried and the improvement in performance. That correlation is significant at a level of 99%. That indicates that queries that have a low baseline performance tend to improve more from the RF process and might explain why query 4 had a much bigger improvement than the others. This finding is in line with Salton and Buckley's [Salton 1997] conclusions when experimenting with several RF methods on different test collections. They found

that collections that perform relatively poorly in an initial retrieval operation can be improved more significantly in a feedback search.

Topic	Baseline	Average	Change
1	46.33	65.17	+41%
2	26.65	12.87	-52%
3	10.64	12.56	+18%
4	5.99	13.66	+128%
5	16.61	20.21	+22%
6	46.94	47.92	+2%

Table 5.15 – Change in average precision (by topic)

This improvement confirms Dumais' [Dumais 1991] results when applying RF to a monolingual system. In comparison, the overall improvement achieved here was much smaller, however, the conditions for the two experiments were different. Her experiments did not employ humans to assess relevance and were done in different test collections. The tests done here built upon Dumais' by employing a CLIR scenario and users from a different language to assess the relevance of documents.

5.6.2 Machine Translation is as Effective as Hand Translation

Participants assessed machine translated documents with the same accuracy they assessed hand translated documents. Performance was equivalent in terms of false alarm, relevant missed, overlap, and confidence in the judgements made.

The performance of the Portuguese participants from both groups (poor English skills and bilingual) judging hand and automatically translated documents is equivalent to the performance of native English speakers judging English documents. That confirms the conclusions of the main experiment and indicates that MT is as effective as hand translation in aiding users to assess relevance. Those results imply that there is no advantage in judging relevance on hand translated documents despite the extra time and cost incurred in generating them.

These findings are consistent with the results achieved by Bathie and Sanderson [Bathie 2002] who used smaller sample of 8 subjects (all native English speakers). They found no significant difference between the judgements made on texts automatically translated from French and texts originally written in English, though they used a different experimental design.

In fact, the participants' performance was actually better with the machine-translated texts. That seems unreasonable, but it can be attributed to personal differences in judging relevance. The validity of relevance assessments is discussed in detail in section 6.1.

The superior performance of machine translated texts in comparison to the judgements made on the original texts happened despite the many translation errors and awkward grammar, which led to several complaints from the participants. Table 5.16 shows some translation errors committed by SYSTRAN.

	English Term	SYSTRAN Translation	Meaning of Translation	Correct Translation
1.	the reaction was skeptical	a reação era skeptical	(meaningless)	a reação era cética
2.	Michael Jackson's fans	Os ventiladores de Michael Jackson	Michael Jackson's ventilators	Os fãs de Michael Jackson
3.	The Princess of Rock	A princesa da rocha	The Princess of the stone	A Princesa do Rock
4.	When the man was late picking Tikriti up	quando o homem era colheita atrasada Tikriti acima	when the man was late harvest Tikriti above	quando o homem atrasou-se para buscar Tikriti
5.	Soviet rule	régua soviética	Soviet ruler (tool)	domínio soviético
6.	Kristen Pfaff found dead	Kristen Pfaff encontrado inoperante	Kristen Pfaff found inoperative	Kristen Pfaff encontrada morta
7.	United States and Britain	estados e Grã Bretanha unidos	states and Britain united	Estados Unidos e Grã-Bretanha
8.	market share	parte de mercado	meaningless	fatia de mercado
9.	make out	fazer para fora	make outside	perceber
10.	New York	iork novo	(meaningless)	Nova Iorque

Table 5.16 – Translation Errors

The mistakes made by MT systems are also discussed by Qu et al. [Qu 2000]. They used SYSTRAN as well, but the language pair was English-French. They have classified the errors according to their type.

- Missing translation, see error 1.
- Unnecessary translation, i.e. the term was replaced by a literal translation when it should remain in the original language, see error 2.
- Wrong sense disambiguation, see errors 4, 5 and 6.
- Wrong phrase, i.e. combining words that do not belong to the same noun phrase, see error 7.
- Word-by-word translation of a multiword term, see error 8

The errors found here cover all those types, and a couple more can be added: Phrasal verb and proper name recognition. Those are shown in Table 5.16 on rows 9 and 10 respectively.

It is difficult to determine the effect that those errors had in the understanding of the texts. However, given that participants' performance with MT texts equals the performance with hand translated texts, there is an indication that those errors did not significantly impair the judgement process. Nevertheless, such errors may affect other tasks such as summarising and capturing more specific information (see section for 6.1 a discussion).

5.6.3 Misjudged Documents have a Moderate Effect

It was expected that the more mistakes a participant made, the less the improvement, resulting in a strong negative correlation between the two variables. However, that assumption turned out to be false. The official judgements did not provide optimal performance (see Table 5.10). The participant who made the least mistakes and achieved the highest overlap was not the one who attained the greatest performance improvement. Similarly, the participant who made the most mistakes was not the one who attained the largest decrease in performance.

The different behaviour of topics in response to the judgement mistakes indicates that the factors that affect the performance of RF vary from one topic to another.

This finding seems analogous to the fact that different queries are better solved by different IR systems. In other words, if a system achieved a high precision in one query, that does not determine it will achieve a good result with another topic. Mandl and Womser-Hacker [Mandl 2003a] have also shown this when evaluating several CLEF runs. They observed a high standard deviation for the performance of the topics and a high standard deviation for the performance of each run. They concluded that no run performed well in all topics. Presently, there are no means for determining which type of topics will do better with which type of IR system.

In conclusion, this study was able to establish that the effects of misjudged documents are different for each topic. It was shown that the main source of impact on the change in performance produced by RF is the topics, and not the users or the systems. However, the characteristics of the topics that determine the relationship between change in performance and the misjudged documents remain unclear. An analysis of the characteristics of the topics would require a larger number of queries than the one used in this experiment.

5.6.4 Difficulty of Task and Accuracy in Judgement

Some research found in the literature analyses the relationship between the perceived difficulty of the query topics and the IR system's performance. Voorhees and Harman [Voorhees 1997] asked topic creators and Natural Language Processing specialists to predict the difficulty of the topics. The participants were asked to classify each topic as easy, medium or difficult. When TREC results had been submitted they were able to assess the accuracy of the predictions. Their results show that the largest correlation between a prediction and the true performance was 0.25. They concluded that no participant was able to accurately predict the true difficulty and there was little agreement among participants. Mandl and Womser-Hacker [Mandl 2003b] performed a similar analysis on CLEF topics. They asked the topic creators to judge the difficulty of the topics and compared it against the performance of the systems. They found a

weak correlation of 0.14. Eguchi, Kuriyama and Kando [Eguchi 2002] analysed NTCIR topics reaching the same conclusion.

A weak correlation of 0.18 was found between the difficulty of the task and the improvement in precision in this experiment. Although measuring different things, the conclusion reached here is in line with the findings mentioned above. People are not able to tell with accuracy what topics will be more difficult for an IR system, in the same way their performance in judging documents was only marginally influenced by how difficult they considered the task to be.

5.6.5 Proper Names help the RF process

It seems logical that the presence of proper names from the queries such as Nirvana or Solzhenitsyn in the documents would provide the user with great help for the judgements. This was expected to aid the assessment of the English documents in particular. However, the results do not confirm that assumption, as even with the presence of such keywords, most users were not able to accurately judge documents. A possible reason for it is that the participants did not rely on the presence of such clues alone, when they could not understand the context in which such proper names were used.

In a recent study of the CLEF topics, Mandl and Womser-Hacker [Mandl 2003b] found that the use of proper names helps CLIR systems. Analysing CLEF results, they concluded that the more proper names a topic contains, the better the retrieval results are. However, the same analysis on the topics chosen for this experiment disagrees with Mandl & Womser-Hacker's conclusion. The topic with the lowest average precision (topic 4) is the one that contains the most proper names. They found a positive correlation of 0.43 between the number of proper names and the average precision of the topic. The figure for the topics used here is -0.23. However, when analysing the relationship between the number of proper names and the improvement in performance achieved through RF a positive correlation of 0.35 was found. It can be concluded that the presence of

proper names did not help the participants in making judgements nor did it have a positive effect in the performance of the baseline run, but it nevertheless had a positive impact on the feedback process.

5.6.6 Participants Feedback

The participants were asked to rate a system that accepts queries in one language and searches for documents in other languages in terms of its usefulness. 92% of the searchers said such a system would be very useful.

They were also asked in which language they preferred the results to be shown. 88% of the participants replied that they preferred the results translated into their language and 12% said they would rather receive the documents in the original language. Those answers confirm the need for a working CLIR system.

5.7 Summary of Chapter

This chapter reported experiments to evaluate relevance feedback on a Cross-Language Information Retrieval system. Portuguese speakers were asked to judge the relevance of some documents returned in response to an initial query. The 27 participants recruited have assessed English documents, documents hand-translated to Portuguese, and documents automatically translated to Portuguese. The accuracy of such judgements was evaluated by comparing them to the official relevance assessments provided by CLEF. In addition the relationship between accuracy in judgement and the performance of RF was studied. The main findings are summarised below:

- Less than half (44%) of the participants were able to assess English documents.
- Machine Translation can indeed aid searchers in making relevance assessments, despite producing documents that are awkward to read.

Participants judged machine translation documents with the same accuracy they judged hand-translated documents.

- There is a moderate negative correlation between the number of misjudged documents and the improvement that Relevance Feedback can provide.
- The factors that impact the change in performance vary greatly from one topic to another.
- The number of proper names present in the topic is positively correlated with the improvement in performance.
- The presence of proper names in the topics did not help users in judging relevance.
- No relationship was found between the change in performance and the difficulty of the topics or the confidence in the assessments or the knowledge of the subject.
- Most participants would consider a CLIR system as very useful and would like the results translated into their native language.

6 DISCUSSION

This Chapter presents a broader look at the findings presented in Chapter 5. The work done is examined in the larger context of the field. The discussion topics were divided into 5 sections. Section 6.1 discusses the applications of machine translation; Section 6.2 addresses common criticisms to the IR experimental evaluation; Section 6.3 debates the use of RF in the real world; Section 6.4 discusses some algorithmic considerations regarding LSI; and Section 6.5 analyses whether it is possible to generalise from the results obtained from the experiments.

6.1 What is MT Good For?

It is widely known that MT systems commit several mistakes. However, the extent to which that affects the user in comprehending the text has not been measured yet. The experiment reported in Chapter 5 showed that MT is as good as hand translation for the purpose of assessing relevance. Nonetheless, whether people can actually extract some useful information from the machine translated texts (i.e. effectively make use of the information contained in the text) was outside the scope of the experiment and remains an open question.

Taylor and White [Taylor 1998] have established a set of tasks that can be used to assess MT systems. The idea is to evaluate MT systems in terms of tasks for

which their output can be successfully used. The tasks proposed by the paper are presented in Table 6.1.

Task	Description
Publishing	Produce a technically correct document in fluent English
Gisting	Produce a summary of the document
Extraction	For documents of interest, capture specified key information
Triage	For documents determined to be of interest, rank by importance
Detection	Find documents of interest
Filtering	Discard irrelevant documents

Table 6.1 – Ranking of Text-Handling Tasks [Taylor 1998]

According to that ranking, the task performed by the participants of the experiment described in Chapter 5 is a combination of filtering and detection. Those are at the bottom of the list, assuming a minimal level of performance from the MT system. Experiments applying users to analyse whether MT can be employed for the other tasks have not been found.

6.2 Addressing Common Criticisms of the Methodology

This section aims at addressing the most common criticisms to the IR experimental evaluation which concern the design of the experiments carried out for this work. The aspects discussed are: validity of relevance assessments, relevance assessments for a cross-linguistic environment, and the use of ready query topics.

6.2.1 Validity of Relevance Assessments

Several studies criticise the traditional IR experimental evaluation, which uses relevance as its basis. Some of the complaints are regarding the simplistic view that assigns binary values to relevance, and some concern the suspected volatile nature of relevance.

Harter [Harter 1996] affirms that he does not believe that there exists a valid interpretation of the meaning of the results of retrieval testing that are based on

fixed, unchanging relevance judgements. This statement is based on the fact that even for carefully constructed test collections such as the Cranfield in which all documents have been judged, subsequent studies [Harter 1971; Swanson 1971] have found missed relevant documents and disagreement among judges.

Schamber [Schamber 1994] affirms that relevance judgements can vary from one assessor to another and even for the same assessor at different times. That claim disagrees with the findings of Resnick and Savage [Resnick 1964] who tested the consistency of human judgements of relevance for 60 subjects. Each subject was asked to judge 34 items in terms of relevance to the subject's personal interests. The items were citations, abstracts or the full text of the documents. One month after the experiment, the same subjects were asked to judge the same items again. The results show that 90% of the items were judged in the same way.

Some research on the nature of relevance, including Park's [Park 1993] which shows the variability of relevance judgements, have presented users with bibliographic citations and observed the criteria they used for following or rejecting a citation. In addition, no definition of relevance was given to the participants. That experimental setting is far more prone to variation and subject to user bias than a scenario in which the user is presented with the full text of the documents and is given a definition of what relevance means in the context of the experiment. Thus, some of the issues identified by those studies do not apply to the experiments carried out here.

Judgements are affected by several factors, some to do with the user and some to do with the query topic or the documents retrieved in response to the query. A number of factors are listed in Table 6.2.

Schamber [Schamber 1994] identified as many as 80 factors that have been suggested in the literature to affect relevance judgements. Nevertheless, fewer than half of those factors have actually been tested by an early experiment by Cuadra and Katter [Cuadra 1967].

<ul style="list-style-type: none">• Type of Document<ul style="list-style-type: none">• Full text• Abstract• Title• Stylist Characteristics of Documents and Queries• Difficulty of Topic• Time pressures• Instructions given to judges• Characteristics of judges<ul style="list-style-type: none">• Knowledge of topic• Experience• Intelligence

Table 6.2 – Some factors that may affect the judgement of relevance

Several studies mention factors that affect relevance judgements [Harter 1996; Park 1993; Schamber 1994], but very few have actually investigated the degree to which the variation of such judgements affects the evaluation performance of retrieval systems. This issue has been addressed by Burgin [Burgin 1992] and Voorhees [Voorhees 1998]. Burgin has done an investigation based on the Cystic Fibrosis collection which is composed by 1239 papers, 100 queries, and four sets of relevance judgements (three from subject experts and one from a medical bibliographer). The existence of these various sets of relevance enabled the investigation. In addition, six different document representations were used (including: abstracts only, title only, title+abstract+heading) The overlap between different sets of judgements ranged between 36% and 59%. The results of retrieval evaluation state that there was no noticeable difference due to variations in relevance judgements, i.e. the difference in retrieval performance with different sets of judgements was less than 5%. Lesk and Salton [Lesk 1968] provide an explanation to the lack of impact of the judgement variations on the retrieval performance. They state that most disagreements among judges concern borderline cases and not the highly relevant documents. Such borderline cases would normally receive a low position on the ranked list of retrieved documents, thus their effect on recall and precision may be expected to be negligible.

Voorhees [Voorhees 1998] analysed two aspects with response to variations in the relevance judgements: the change in average precision, and the change in

system rankings (i.e. to establish whether system A is better than system B). The investigation was done using the TREC collection. Several different conditions were tested, including judgements made by the author of the query vs. judgements made by non-authors; judgements made by a single assessor vs. judgements made by a group; and judgements made by people with different backgrounds. The overlap between different assessors ranged between 32% and 49%. The conclusions are that the average precision of the runs was affected by the changes in judgements. In some occasions the difference in mean average precision was greater than 10%, which is considered significant. However, system rankings was stable as every time a run would score higher (or lower) with one set of judgements, so would the other runs.

In summary, the history of the criticism of the IR evaluation method is almost as old as the field itself. Several critics of the approach have pointed to factors that may affect judgements of relevance. On the other hand, few experiments have investigated the extent to which disagreements among judges actually affect the evaluation process. Respectable IR researchers such as Salton and Lesk have regarded the effect of judgement variations as negligible. One possible explanation for such disagreement amongst researchers is that different tasks will produce different degrees of variability in judgements. In conclusion, relevance assessments are variable and this variation may affect the systems score. However, that effect is limited as relevance is far from being a random trend. In addition, the factors thought to influence the judgements can be minimised in an experimental setting, e.g. participants can be given a standardised definition of relevance, documents can be presented using a uniform style, time constraints can be kept constant, etc. Finally, IR systems need to be evaluated using some criteria, and despite the criticisms, relevance is still the norm.

6.2.2 CLEF Topics and Relevance Assessments

If there are disagreements on the relevance of documents for judges that share the same background, it seems plausible that this problem will escalate in the presence of another language and, by consequence, another culture. However, the results from the experiment do not show that. The performance of Portuguese participants judging the relevance of translated documents is comparable to the performance of the English assessors. Furthermore, the level of agreement achieved by them is equivalent to the level of agreement found by Voorhees [Voorhees 1998] when analysing the judgements made by people from the same background.

One possible reason for this homogeneity is that CLEF topics are created with a special care to avoid cultural bias that could influence the semantics of the topic in different languages. A thorough translation check of all translated topics, made by professional translators assures that all versions include the same semantics. Aspects of multilingual topic creation and relevance assessments for the CLEF campaign are discussed in greater detail by Kluck [Kluck 2002].

6.2.3 The Use of Ready Query Topics

There is some criticism of the use of ready queries for user experiments. Ready queries are not considered to be real “information needs” [S. E. Robertson 1992]. That criticism is mostly directed to early IR collections which used elaborated queries giving systems a good basis for retrieval. Present collections use shorter topics that resemble real user queries.

It seems plausible that users would be able to judge with more accuracy topics that they are familiar with. In other words, documents retrieved in response to a real user query would be judged with greater precision. However, the results from the experiment in Chapter 5 show a weak correlation of 0.27 between the

knowledge of topic and the overlap. Indicating that only 7% of the accuracy in judgement can be explained by the difficulty of the topic.

One possible reason is that CLEF topics deal with real world affairs, such as politics, sports, science and culture. Despite not being created by the participants themselves they still simulate a real information need, and that is enough for the judgement process. In addition, the use of real user queries in this experimental setting would not be feasible as the relevance assessments should be known in advance in order to allow for the calculation of the measurements of retrieval performance and judgement accuracy.

6.3 RF in the Real World

Despite being widely applied to IR, and having its advantages demonstrated, it is not clear whether users would really be willing to use RF in practice. In a study of real user queries on the web Jansen, Spink and Saracevic [Jansen 1999] analysed transaction logs from over 51,000 queries posed by 18,000 users to Excite¹⁰, a popular web search engine. They found that only 5% of the queries used relevance feedback. In order to use the RF option on that search engine the user needs only to click on a “More Like This” hyperlink that implements the RF mechanism.

The authors [Jansen 1999] also try to establish whether RF runs were successful. That was made based on some assumptions: (i) if the user quit after the feedback run, that was counted as success; (ii) if the user returned to the exact previous query after RF, that was counted as failure; (iii) if after the RF process the user returned to a similar query, that was classified as a less than successful since RF could have provided some terms suggestions; and (iv) if the user returned with a totally different query that was considered browsing. Analysing the data, 63% of

¹⁰ <http://www.excite.com>

the feedback runs were considered successful, 17% were considered less than successful; 16% were considered failures; and 4% fell on the browsing category. The study does not speculate on possible causes for this lack of utilisation. One possible reason is that most users of Internet search engines are novice to computer searching, and are not aware of the benefits of the procedure.

6.4 Algorithmic Considerations

This section discusses the feasibility of applying LSI to large scale IR, and the benefits of incorporating negative feedback to LSI's RF method.

6.4.1 LSI and Large Scale IR

LSI dramatically reduces the number of dimensions of the document space by only keeping the most important concepts. Those concepts are far fewer than the number of words in the document collection. However, the calculation of the SVD, which reduces the number of dimensions is computationally expensive. For the tests done in Chapter 4, the indexing process took as long as 2 days when using 400 dimensions. That makes LSI prohibitive for very large document collections. Another problem is that the vectors generated by LSI are dense, therefore searching time can be long. However, more efficient methods for reducing the number of dimensions are being researched. Hoenkanp [Hoenkamp 2001] proposes using Haar Transform instead of SVD claiming savings in computational cost. Kolda and O'Leary [Kolda 1998] propose applying Semidiscrete Decomposition as an alternative to SVD. They report savings in storage space and in query processing time.

LSI is a relatively new technique for IR. Despite the computational cost of deriving the semantic space, SVD algorithms are becoming faster. Dumais [Dumais 1995] reported that the algorithm used for TREC-3 was much faster than the original implementation. In addition, efficient alternatives to SVD are being

researched. Also, Excite claims to have implemented a more efficient version of LSI [Zakis 1997]. Those facts suggest that LSI may have a future on large scale IR.

6.4.2 Using Negative Feedback

The RF experiments presented in Chapter 5 only made use of positive feedback, which meant that if a participant did not classify any document as relevant, RF could not be performed. That represents a shortcoming since if no documents are found relevant, the query could be modified and moved away from those irrelevant documents.

In the context of the experiment, however, the only cases in which a participant did not select any relevant documents was when he did not understand the contents of the text, i.e. when judging English documents. In that situation, the use of negative feedback would not be useful since the meaning of the lack of positive feedback was different.

Nevertheless, previous research on RF for vector models has shown that negative feedback can also be used to improve performance. That fact justifies the proposal of new RF methods for LSI that incorporate negative feedback.

6.5 Generalising from the Results

Since participants were selected from the population of interest, and the sample and the number of queries was significant it is possible to reasonably generalise beyond the individuals tested, to the population from which they were drawn. The experiments presented in Chapter 5 examine mainly two aspects: participant's ability in assessing translated documents and the relationship between accuracy in judgement and RF improvement.

The findings concerning the first aspect can be reasonably generalised. It is logical to think that if a different language pair was applied using the same experimental setting, the same results would be achieved. That implies MT can

be successfully used to serve as a basis for relevance judgements in other languages.

The findings for the second aspect, however, are harder to generalise. The results have shown a great disparity among topics. Therefore, it is likely that the choice of topics will impact the findings. In order to be able to generalise from the results, a much larger set of topics needs to be used. In addition, it is possible that the effect that misjudged documents have on the performance of RF is influenced by the IR system.

7 CONTRIBUTIONS AND FUTURE WORK

This is the concluding chapter of the thesis, in which the contributions are presented and possibilities of future work are discussed.

7.1 Contributions of the Thesis

This thesis presented four main contributions to the field: CLIR was extended to another language pair; an effective stemming algorithm for Portuguese was developed; several LSI-CLIR experiments were carried out; the use of translated texts for relevance assessment was analysed; and the relationship between accuracy in judgement and performance improvement was studied.

7.1.1 CLIR was extended to a new language pair

Portuguese is the fifth biggest language in number of native speakers. It is estimated that there are more than 8 million Portuguese speakers accessing the Internet. The problem is that less than 1% of the information available is written in Portuguese, the vast majority being in English. This situation leaves the Portuguese population with a very limited source for information. The work described here represents a first step on Portuguese-English CLIR, which aims at making more information available to the Portuguese speaking population.

7.1.2 Development of a Portuguese stemming algorithm

Despite being a major world language, Portuguese is extremely under-represented compared to languages such as French, Spanish and Italian. The development of a Portuguese stemming algorithm represented an effort towards increasing the use of Portuguese for IR experiments. Since its development, several requests were made for the algorithm and it is in use in fifteen different institutions. The algorithm has been shown to be effective and improves the performance of a CLIR system.

7.1.3 Running of several LSI - CLIR experiments

Chapter 4 describes several CLIR experiments done using LSI. They involve different test collections and a new language-pair. Those tests provided a demonstration that MT can be effectively used to simulate parallel corpora. This finding is consistent with Lavrenko's [Lavrenko 2002]. In addition, a comparison of the performance between monolingual and bilingual runs was made. Furthermore, the validity of stemming was analysed. The results of such experiments contribute data points to the CLIR theory.

7.1.4 Analysis of translated texts for relevance assessment

A user experiment was carried out considering RF on a CLIR environment. Portuguese speakers representing likely users of a CLIR system were selected and asked to judge a set of documents for relevance. The documents were either: the original English text, a hand translation or an automatic translation produced by SYSTRAN. The results show that MT is as effective as hand translation in aiding participants to assess relevance. The performance of Portuguese participants assessing translated texts is comparable to the performance of native English speakers judging texts in English.

7.1.5 Study of the relationship between accuracy in judgement and performance improvement

The relationship between the accuracy in judgements and the improvement in performance provided by RF remained unexplored. No research found in the literature deals with that aspect. The experiment carried out by this thesis has shown that the impact of misjudged documents varies immensely from one query topic to another and in some cases even contributes to performance improvements.

7.2 Future Work

From the work developed in this thesis, a number of options for future work emerged. These are discussed in the next sections.

7.2.1 Variations on RF Experiment

The results of the RF experiment can be generalised to other languages. In order to confirm that, another language-pair could be tested, applying the same experimental design.

The RF experiments have been done using LSI. It is possible that different IR systems will produce different results. It would be interesting to compare the effects that misjudged documents have on the performance of RF on different IR systems. Similarly the experiments could be repeated with different topics to allow for comparisons.

7.2.2 Analyse the characteristics of the topics

The experiments done in this thesis have shown that the impact of misjudged documents on RF varies greatly for different query topics. It would be desirable to have a mechanism for specifying which characteristics of the topics determine the effect on performance. This analysis would require an experiment containing

a large number of topics. Relevance judgements do not necessarily need to be obtained from users, as they could be automatically generated.

7.2.3 Propose a method for Negative Feedback

Presently, RF methods for LSI only make use of positive judgements. A possible area for future work would be proposing and testing RF methods that deal with negative feedback, as it has been considered useful applied to other vector methods.

7.2.4 Improving the Portuguese stemmer

The Portuguese stemmer proposed in this thesis has performed better than the Portuguese version of the Porter stemmer in several tests. It would be interesting, however, to combine some features of the Porter stemmer into the RSLP stemmer aiming at improving accuracy. One possibility is to use Porter's measure for length of words which is based on lists of vowels and consonants to help decide whether or not it is wise to take off a suffix. RSLP uses absolute numbers to make that decision (minimum stem size) which intuitively seems to be less effective.

7.2.5 Use comparable corpora for training the IR system

The experiments done here used parallel corpora to derive LSI's vector-space. The problem is parallel corpora are scarce, and were not available for the language-pair in question. Therefore, MT was used to generate a sample of dual-language documents. Comparable corpora, however, seems to be more readily available. One possible extension to LSI's application for CLIR would be the development of methods that allow the use of comparable, rather than parallel corpora to derive the vector-space. Recently, a corpus containing news from a Brazilian newspaper (Folha de São Paulo) has been released. The year is the same for the LA Times collection used here, therefore it is likely that some articles deal with similar subjects allowing the creation of a comparable corpus.

7.3 Final Word

CLIR is an interesting multidisciplinary subject that has been increasingly gaining more attention from the research community. Despite recent advances and new developments, there are still many aspects to be explored.

The work presented here represents a contribution in many ways; by testing a new language pair; by developing linguistic resources; by showing that MT can be effectively used to simulate parallel corpora; by analysing the use of MT for relevance assessments; and by studying the effects of the errors in judgement.

There are also several possibilities for future developments as it is possible to build up on the work completed here.

REFERENCES

[Adriani 2000]

M. Adriani. (2000). Using Statistical Term Similarity for Sense Disambiguation In Cross-Language Information Retrieval. *Information Retrieval*, 2(1), 67-78.

[Baeza-Yates 1999]

R. Baeza-Yates & B. Ribeiro-Neto. (1999). *Modern Information Retrieval*. Harlow, England: ACM Press.

[Ballesteros 2000]

L. Ballesteros. (2000). Cross-Language Retrieval via Transitive Translation. In W. B. Croft (Ed.), *Advances in Information Retrieval - Recent Research from the Center for Intelligent Information Retrieval* (pp. 306): Kluwer Academic Publishers.

[Ballesteros 1997]

L. Ballesteros & W. B. Croft. (1997). Phrasal Translation and Query Expansion Techniques for Cross-Language Information Retrieval. In *AAAI Spring Symposium on Text and Speech Retrieval*. Stanford, California: AAAI Press.

[Ballesteros 1998]

L. Ballesteros & W. B. Croft. (1998). Statistical Methods for Cross-Language Information Retrieval. In G. Grefenstette (Ed.), *Cross-Language Information Retrieval* (pp. 23-40). Boston: Kluwer Academic Publishers.

[Bathie 2002]

Z. Bathie & M. Sanderson. (2002). iCLEF at Sheffield. In C. Peters, M. Braschler, J. Gonzalo & M. Kluck (Eds.), *Evaluation of Cross-Language Information Retrieval Systems. Second Workshop of the Cross-Language Evaluation Forum, CLEF 2001. Lecture Notes in Computing Science 2406* (pp. 332-335): Springer.

[Biron 1995]

P. V. Biron & D. H. Kraft. (1995). New Methods for Relevance Feedback: Improving Information Retrieval Performance. In *ACM Symposium on Applied Computing*. Nashville, Tennessee.

[Braschler 2000]

M. Braschler, D. Harman, M. Hess, M. Kluck, C. Peters & P. Schäuble. (2000). Evaluation of Systems for Cross-Language Information Retrieval. In *Second International Conference on Language Resources & Evaluation (LREC)*. Athens, Greece.

[Braschler 1998]

M. Braschler, C. Peters, E. Picchi & P. Schäuble. (1998). Cross-Language Web Querying: The EuroSearch Approach. In C. Nikolau & C. Stephanidis (Eds.), *Research and Advanced Technology for Digital Libraries: Second European Conference (ECDL)*. Heraklion, Crete, Greece.

[Braschler 1999]

M. Braschler, C. Peters & P. Schäuble. (1999). Cross-Language Information Retrieval Track Overview. In *Text Retrieval Conference (TREC-8)*. Gaithersburg, Maryland: NIST.

[Brisaboa 2002]

N. R. Brisaboa, C. Callón, J. R. López, A. S. Places & G. Sanmartin. (2002). Stemming Galician Texts. In A. F. Leander & A. L. Oliveira (Eds.), *9th International Symposium on String Processing and Information Retrieval (SPIRE)Lisbon, Portugal - LNCS 2476*: Springer.

[Burgin 1992]

R. Burgin. (1992). Variations in Relevance Judgements and the Evaluation of Retrieval Performance. *Information Processing & Management*, 25(5), 619-627.

[Chang 1971]

Y. K. Chang, C. Cirillo & J. Ranzon. (1971). Evaluation of Feedback Retrieval Using Modified Freezing, Residual Collection, and Test and Control Groups. In G. Salton (Ed.), *The SMART Retrieval System. Experiments in Automatic Document Processing* (pp. 335-370). Englewood Cliffs, New Jersey: Prentice Hall.

[Chen 2003]

A. Chen. (2003). Cross-Language Experiments at CLEF-2002. In C. Peters, M. Braschler, J. Gonzalo & M. Kluck (Eds.), *Advances in Cross-Language Information Retrieval - Third Workshop of the Cross-Language Evaluation Forum, CLEF 2002 (LNCS 2785)* (pp. 5-20): Springer.

[CLEF]

CLEF.Cross-Language Evaluation Forum. Retrieved on 28/10/2003, from <http://www.clef-campaign.org>

[Crestani 2000]

F. Crestani. (2000). Neural Relevance Feedback for Information Retrieval. In B. Bouchon-Meunier, R. R. Yager & L. A. Zadeh (Eds.), *Uncertainty in Intelligent in Information Systems*. Singapore: World Scientific.

[Croft 1983]

W. B. Croft. (1983). Experiments with Representation in a Document Retrieval System. *Information Technology: Research and Development*, 2(1), 1-21.

[Cuadra 1967]

C. A. Cuadra & R. V. Katter. (1967). Opening the Black Box of Relevance. *Journal of Documentation*, 23, 291-303.

[Cunha 1985]

C. Cunha & L. Lindley-Cintra. (1985). *Nova Gramática do Português Contemporâneo*. Rio de Janeiro: Nova Fronteira (in Portuguese).

[Davis 1998]

M. W. Davis. (1998). On the Effective use of Large Parallel Corpora in Cross-Language Information Retrieval. In G. Grefenstette (Ed.), *Cross-Language Information Retrieval* (pp. 11-21). Boston: Kluwer Academic Publishers.

[Deerwester 1990]

S. Deerwester, S. Dumais, G. Furnas, T. Landauer & R. Harshman. (1990). Indexing by Latent Semantic Analysis. *Journal of the American Society for Information Science*, 41(6), 1-13.

[Dillon 1980]

M. Dillon & J. Desper. (1980). Automatic Relevance Feedback In Boolean Retrieval System. *Journal of Documentation*, 36, 197-208.

[Drucker 2001]

H. Drucker, B. Shahary & D. Gibbon. (2001). Relevance Feedback using Support Vector Machines. In *18th International Conference on Machine Learning (ICML)*.

[Dumais 1991]

S. Dumais. (1991). Improving Retrieval of Information from External Sources. *Behaviour Research Methods, Instruments and Computers*, 23(2), 229-269.

[Dumais 1995]

S. Dumais. (1995). Using LSI for Information Filtering: TREC-3 experiments. In D. Harman (Ed.), *Third Text REtrieval Conference (TREC-3)*: NIST.

[Dumais 1992]

S. Dumais & J. Nielsen. (1992). Automating the Assignment of Submitted Manuscripts to Reviewers. In *15th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. Denmark.

[Efthimiadis 2000]

E. Efthimiadis. (2000). Interactive Query Expansion: A User-based Evaluation in a Relevance Feedback Environment. *Journal of the American Society for Information Science*, 51(11), 989-1003.

[Eguchi 2002]

K. Eguchi, K. Kuriyama & N. Kando. (2002). Sensitivity of IR Systems Evaluation to Topic Difficulty. In *3rd International Conference on Language Resources and Evaluation (LREC 2002)* (Vol. 2, pp. 585-589). Las Palmas de Gran Canaria, Spain.

[EMIR-Consortium 1994]

EMIR-Consortium. (1994). *Final Report of the EMIR project 5312*. Brussels: Commission of the European Union.

[Erbach 1997]

G. Erbach, G. Neumann & H. Uszkoreit. (1997). Multilingual Indexing, Navigation and Editing Extensions for the World Wide Web. In *Third DELOS Workshop on Cross-Language Information Retrieval*. Zurich, Switzerland.

[Ethnologue]

Ethnologue. Retrieved on 18/07/2003, from <http://www.ethnologue.com>

[Ferber 1997]

R. Ferber. (1997). Automated Indexing with Thesaurus Descriptors: A Co-occurrence Based Approach to Multilingual Retrieval. In C. Peters & C. Thanos (Eds.), *Research and Advanced Technology for Digital Libraries: First European Conference*. Pisa, Italy.

[Figuerola 2000]

C. G. Figuerola, J. L. A. Berrocal, A. F. Z. Rodriguez & R. G. Diaz. (2000). Retrieval of Spanish-English Information by Means of a Standard Automatic Translation System. In C. Peters (Ed.), *Cross-Language Information Retrieval and Evaluation, Workshop of the Cross-Language Evaluation Forum CLEF 2000, Lisbon, Portugal (LNCS 2069)*: Springer.

[Fluhr 1999]

C. Fluhr. (1999). Crosslingual Interrogation of Multilingual Catalogs. In *Third European Conference on Research and Advanced Technology for Digital Libraries (ECDL)*. Paris, France.

[Fluhr 1997]

C. Fluhr, D. Schmit, F. Elkateb, P. Ortet & K. Gurtner. (1997). Multilingual Database and Crosslingual Interrogation in a Real Internet Application. In *AAAI Spring Symposium on Cross-Language Text and Speech Retrieval* (pp. 32-36). Stanford, California: AAAI Press.

[Foltz 1990]

P. W. Foltz. (1990). Using Latent Semantic Indexing for Information Filtering. In *ACM SIGOIS Bulletin - Proceedings of the Conference on Office Information Systems*.

[Foltz 1992]

P. W. Foltz & S. Dumais. (1992). Personalized Information Delivery: An Analysis of Information Filtering Methods. *Communications of the ACM*, 35(12), 51-60.

[Forsythe 1977]

G. E. Forsythe, M. A. Malcolm & C. B. Moler. (1977). Least Squares and the Singular Value Decomposition. In *Computer Methods for Mathematical Computations*. Englewood Cliffs, New Jersey: Prentice Hall.

[Frakes 1992]

W. B. Frakes & R. Baeza-Yates. (1992). *Information Retrieval - Data Structures and Algorithms*. London, UK: Prentice Hall.

[Franz 2002]

M. Franz & J. S. McCarley. (2002). Arabic Information Retrieval at IBM. In E. Voorhees & L. P. Buckland (Eds.), *The Eleventh Text Retrieval Conference - TREC2002*. Gaithersburg, Maryland: NIST.

[Frederking 1997]

R. Frederking, T. Mitamura, E. Nyberg & J. Carbonell. (1997). Translingual Information Access. In *AAAI Spring Symposium on Cross-Language Text and Speech Retrieval* (pp. 40-48). Stanford, California: AAAI Press.

[Gachot 1998]

D. A. Gachot, E. Lange & J. Yang. (1998). The SYSTRAN NLP Browser: An Application of Machine Translation Technology in Cross-Language Information Retrieval. In G. Grefenstette (Ed.), *Cross-language Information Retrieval*. Boston: Kluwer Academic Publishers.

[Gao 2000]

J. Gao, J. Nie, J. Zhang, E. Xun, Y. Su, M. Zhou, et al. (2000). TREC-9 CLIR Experiments at MSRCN. In E. Voorhees & D. Harman (Eds.), *The 9th Text Retrieval Conference - TREC-9*. Gaithersburg, Maryland: NIST.

[Gey 1999]

F. C. Gey & H. Jiang. (1999). English-German Cross-Language Retrieval for the GIRT Collection - Exploiting a Multilingual Thesaurus. In *Eighth Text Retrieval Conference (TREC-8)*. Gaithersburg, Maryland.

[Gilarranz 1997]

J. Gilarranz, J. Gonzalo & F. Verdejo. (1997). An Approach to conceptual Text Retrieval using the EuroWordNet Multilingual Semantic Database. In *AAAI Spring Symposium on Cross-Language Text and Speech Retrieval* (pp. 49-55). Stanford, California: AAAI Press.

[Gollins 2000]

T. M. Gollins. (2000). *Dictionary Based Transitive Cross-Language Information Retrieval using Lexical Triangulation*. PhD Thesis, University of Sheffield, Sheffield, UK.

[Grefenstette 1998]

G. Grefenstette. (1998). *Cross-Language Information Retrieval*. Boston: Kluwer Academic Publishers.

[Harman 1991]

D. Harman. (1991). How Effective is Suffixing? *Journal of the American Society for Information Science*, 42(1), 7-15.

[Harman 1992]

D. Harman. (1992). Relevance Feedback Revisited. In *15th ACM SIGIR Conference on Research and Development in Information Retrieval*. Denmark.

[Harter 1971]

S. P. Harter. (1971). The Cranfield II Relevance Assessments. *Library Quarterly*, 41, 229-243.

[Harter 1996]

S. P. Harter. (1996). Variations in Relevance Assessments and the Measurement of Retrieval Effectiveness. *Journal of the American Society for Information Science*, 47(1), 37-49.

[Hiemstra 1999]

D. Hiemstra & F. de Jong. (1999). Disambiguation Strategies for Cross-Language Information Retrieval. In *Third European Conference on Research and Advanced Technology for Digital Libraries*. Paris, France.

[Hiemstra 2000]

D. Hiemstra, W. Kraaij, R. Pohlmann & T. Westerveld. (2000). Twenty-One at CLEF-2000: Translation Resources, Merging Strategies and Relevance Feedback. In C. Peters (Ed.), *Cross-Language Information Retrieval and Evaluation, Workshop of the Cross-Language Evaluation Forum CLEF 2000, Lisbon, Portugal (LNCS 2069)*: Springer.

[Hoenkamp 2001]

E. Hoenkamp. (2001). Unitary Operators for Fast Latent Semantic Indexing. In *24th ACM SIGIR Conference on Research and Development in Information Retrieval*. New Orleans, Louisiana.

[Honrado 2000]

A. Honrado. (2000). A Word Stemming Algorithm for the Spanish Language. In *Proceedings of the String Processing and Information Retrieval Conference (SPIRE)*. A Coruña, Spain.

[Hull 1996]

D. A. Hull. (1996). Stemming Algorithms: A Case Study for Detailed Evaluation. *Journal of the American Society for Information Science*, 47(1), 70-84.

[Hull 1997]

D. A. Hull. (1997). Using Structured Queries for Disambiguation in Cross-Language Information Retrieval. In *AAAI Spring Symposium on Cross-Language Text and Speech Retrieval* (pp. 84-98). Stanford, California: AAAI Press.

[Hull 1996]

D. A. Hull & G. Grefenstette. (1996). Querying Across Languages, a Dictionary-based approach to Multilingual Information Retrieval. In *19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. Zurich, Switzerland.

[Ide 1971]

E. Ide. (1971). New Experiments In Relevance Feedback. In G. Salton (Ed.), *The SMART Retrieval System. Experiments in Automatic Document Processing* (pp. 337-354). Englewood Cliffs, New Jersey: Prentice Hall.

[Internet Software Consortium]

Internet Software Consortium. Retrieved on 18/07/2003, from <http://www.isc.org>

[Jansen 1999]

M. B. J. Jansen, A. Spink & T. Saracevic. (1999). *The use of Relevance Feedback on the web: Implications for web IR system design*. Paper presented at the World Conference on the WWW and Internet, Honolulu, Hawaii.

[Karlgren 2003]

J. Karlgren & P. Hansen. (2003). Cross-language Relevance Assessment and Task Context. In C. Peters, M. Braschler, J. Gonzalo & M. Kluck (Eds.), *Advances in Cross-Language Information Retrieval - Third Workshop of the Cross-Language Evaluation Forum, CLEF 2002 (LNCS 2785)* (pp. 255-259): Springer.

[Kluck 2002]

M. Kluck & Womser-Hacker. (2002). Inside the Evaluation Process of the Cross-Language Evaluation Forum (CLEF): Issues of Multilingual Topic Creation and Multilingual Relevance Assessment. In M. G. Rodriguez & C. P. S. Araujo (Eds.), *Proceedings of the Third International Conference on Language Resources and Evaluation, LREC*. Las Palmas de Gran Canaria, Spain: ELRA.

[Kolda 1998]

T. G. Kolda & D. O'Leary. (1998). A Semidiscrete Matrix Decomposition For Latent Semantic Indexing in Information Retrieval. *ACM transactions on Information Systems (TOIS)*, 16(4), 322-346.

[Kraaij 1994]

W. Kraaij & R. Pohlmann. (1994). Porter's Stemming Algorithm for Dutch. In L. G. M. Noordman & W. A. M. de Vroomen (Eds.), *Informatiewetenschap 1994: Wetenschappelijke bijdragen aan de derde STINFON Conferentie*, . Tilburg, Netherlands.

[Krovetz 1993]

R. Krovetz. (1993). Viweing Morphology as an Inference Process. In *16th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 191-202).

[Lam-Adesina 2001]

A. M. Lam-Adesina & G. J. F. Jones. (2001). Applying Summarization Techniques for Term Selection in Relevance Feedback. In *24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 1-9). New Orleans, Louisiana: ACM.

[Landauer 1990]

T. Landauer & M. Littman. (1990). Automatic Cross-Language Document Retrieval using Latent Semantic Indexing. In *Sixth Annual Conference of the UW Centre for the New Oxford English Dictionary and Text Research* (pp. 31-38). Waterloo, Canada.

[Lavrenko 2002]

V. Lavrenko, M. Choquette & W. B. Croft. (2002). "Cross-Lingual Relevance Models". In *25th ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 175-182). Tampere, Finland: ACM.

[Lemaire 2001]

B. Lemaire & P. Dessus. (2001). A System to Assess the Semantic Content of Student Essays. *Journal of Educational Computing*, 24(3), 305-320.

[Lesk 1968]

M. Lesk & G. Salton. (1968). Relevance Assessments and Retrieval System Evaluation. *Information Storage and Retrieval*, 4, 343-359.

[López-Ostenero 2001]

F. López-Ostenero, J. Gonzalo, A. Peñas & F. Verdejo. (2001). Noun Phrase Translations for Cross-Language Document Selection. In C. Peters (Ed.), *Evaluation of Cross-Language Information Retrieval Systems. Second Workshop of the Cross-Language Evaluation Forum, CLEF 2001. Lecture Notes in Computing Science 2406*: Springer.

[Macambira 1999]

J. R. Macambira. (1999). *A Estrutura Morfo-Sintática do Português*. São Paulo, Brazil: Ed. Pioneira (in Portuguese).

[Mandl 2003a]

T. Mandl & C. Womser-Hacker. (2003a). Linguistic and Statistical Analysis of the CLEF Topics. In C. Peters (Ed.), *Advances in Cross-Language Information Retrieval - Third Workshop of the Cross-Language Evaluation Forum, CLEF 2002 (LNCS 2785)*: Springer.

[Mandl 2003b]

T. Mandl & C. Womser-Hacker. (2003b). Proper Names in the Multilingual CLEF Topic Set. In C. Peters (Ed.), *Working Notes for the CLEF 2003 Workshop*. Trondheim, Norway.

[McNamee 2000]

P. McNamee & J. Mayfield. (2000). A Language-Independent Approach to European Text Retrieval. In C. Peters (Ed.), *Cross-Language Information Retrieval and Evaluation Workshop of Cross-Language Evaluation Forum, CLEF 2000 - Lecture Notes in Computer Science 2069*. Lisbon, Portugal: Springer.

[McNamee 2002]

P. McNamee & J. Mayfield. (2002). JHU/APL Experiments at CLEF: Translation Resources and Score Normalization. In C. Peters, M. Braschler, J. Gonzalo & M. Kluck (Eds.), *Evaluation of Cross-Language Information Retrieval Systems - Second Workshop of the Cross-Language Evaluation Forum, CLEF 2001. Lecture Notes in Computer Science 2406*: Springer.

[McNamee 2003]

P. McNamee & J. Mayfield. (2003). Scalable Multilingual Information Access. In C. Peters, M. Braschler, J. Gonzalo & M. Kluck (Eds.), *Advances in Cross-Language Information Retrieval - Third Workshop of the Cross-Language Evaluation Forum, CLEF 2002 (LNCS 2785)*: Springer.

[Moreira Orengo 2001]

V. Moreira Orengo & C. R. Huyck. (2001). A Stemming Algorithm for the Portuguese Language. In *8th International Symposium on String Processing and Information Retrieval (SPIRE)*. Laguna de San Raphael, Chile.

[Moreira Orengo 2003]

V. Moreira Orengo & C. R. Huyck. (2003). Portuguese-English Cross-Language Information Retrieval Using Latent Semantic Indexing. In M. Kluck (Ed.), *Advances in Cross-Language Information Retrieval - Third Workshop of the Cross-Language Evaluation Forum, CLEF 2002 (LNCS 2785)*. Rome, Italy: Springer.

[Nie 2002]

J. Nie & M. Simard. (2002). Using Statistical Translation Models for Bilingual IR. In C. Peters, M. Braschler, J. Gonzalo & M. Kluck (Eds.), *Evaluation of Cross-Language Information Retrieval Systems - Second Workshop of the Cross-Language Evaluation Forum, CLEF 2001. Lecture Notes in Computer Science 2406*: Springer.

[NTCIR]

NTCIR. Retrieved on 05/11/2003, from <http://research.nii.ac.jp/ntcir-ws4/>

[Oard 1997]

D. W. Oard. (1997). Serving Users in Many Languages - Cross-Language Information Retrieval for Digital Libraries. *D-Lib Magazine*.

[Oard 2001]

D. W. Oard & J. Gonzalo. (2001). The CLEF 2001 Interactive Track. In C. Peters, M. Braschler, J. Gonzalo & M. Kluck (Eds.), *Second Workshop of the Cross-Language Evaluation Forum, Lecture Notes in Computer Science 2406*: Springer.

[Oard 2003a]

D. W. Oard & J. Gonzalo. (2003a). The CLEF 2002 Interactive Track. In C. Peters (Ed.), *Advances in Cross-Language Information Retrieval - Third Workshop of the Cross-Language Evaluation Forum, CLEF 2002 (LNCS 2785)* (pp. 245-254): Springer.

[Oard 2003b]

D. W. Oard & J. Gonzalo. (2003b). The CLEF 2003 Interactive Track. In C. Peters (Ed.), *Working Notes for the CLEF 2003 Workshop*. Trondheim, Norway.

[Paice 1994]

C. D. Paice. (1994). An Evaluation Method for Stemming Algorithms. In W. B. Croft & C. J. van Rijsbergen (Eds.), *17th ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 42-50). Dublin, Ireland: ACM.

[Park 1993]

T. K. Park. (1993). The Nature of Relevance in Information Retrieval. *Library Quarterly*, 63, 318-351.

[Peters 1996]

C. Peters & E. Picchi. (1996). *Cross-Language Information Retrieval: A System for Comparable Corpus Querying*. Paper presented at the Working Notes of the Workshop on Cross-Linguistic Information Retrieval - ACM SIGIR, Zurich, Switzerland.

[Peters 1997]

C. Peters & E. Picchi. (1997). Across Languages, Across Cultures - Issues in Multilinguality and Digital Libraries. *D-Lib Magazine*.

[Peters 2000]

C. Peters & P. Sheridan. (2000). Multilingual Information Access. In M. Agosti, F. Crestani & G. Pasi (Eds.), *Lectures on Information Retrieval*. Berlin, Germany: Springer.

[Pirkola 2001]

A. Pirkola, T. Hedlund, H. Keskusato & K. Järvelin. (2001). Dictionary-Based Cross-Language Information Retrieval: Problems, Methods, and Research Findings. *Information Retrieval*, 4(3), 209-230.

[Porter 1980]

M. F. Porter. (1980). An Algorithm for Suffix Stripping. *Program*, 14(3), 130-137.

[Qu 2000]

Y. Qu, A. N. Eilerman, H. Jin & D. A. Evans. (2000). The Effect of Pseudo Relevance Feedback on MT-Based CLIR. In *In Proceedings of Content-Based Multimedia Information Access: Recherche d'Informations Assistée par Ordinateur (RIAO 2000)*. Paris, France.

[Raghavan 1989]

V. V. Raghavan, G. S. Jung & P. Bollman. (1989). A Critical Investigation of Recall and Precision as Measures of Retrieval System Performance. *ACM Transactions on Office and Information Systems*, 7(3), 205-229.

[Redher 1997]

B. Redher, M. Littman, S. Dumais & T. Landauer. (1997). Automatic 3-Language Cross-Language Information Retrieval with Latent Semantic Indexing. In *Sixth Text Retrieval Conference (TREC-6)*. Gaithesburg, Maryland.

[Redher 1998]

B. Redher, M. E. Schreiner, M. B. W. Wolfe, D. Laham, E. Lange & W. Kintsh. (1998). Using Latent Semantic Analysis to Assess Knowledge. *Discourse Processes*, 25, 337-354.

[Renders 2003]

J. M. Renders, H. Dejan & E. Gaussier. (2003). XRCE participation at CLEF2002. In C. Peters (Ed.), *Advances in Cross-Language Information Retrieval - Third Workshop of the Cross-Language Evaluation Forum, CLEF 2002 (LNCS 2785)*: Springer.

[Resnick 1964]

A. Resnick & T. R. Savage. (1964). The consistency of human judgements of relevance. *American Documentation*, 15, 93-95.

[Robertson 1976]

S. Robertson & K. Sparck Jones. (1976). Relevance Weighting of Search Terms. *Journal of the American Society for Information Science*, 27(3), 129-146.

[Robertson 1992]

S. E. Robertson & M. M. Hancock-Beaulieu. (1992). On the Evaluation of IR Systems. *Information Processing & Management*, 28(4), 457-466.

[Rocchio 1971]

J. J. Rocchio. (1971). Relevance Feedback in Information Retrieval. In G. Salton (Ed.), *The SMART Retrieval System. Experiments in Automatic Document Processing* (pp. 313-323). Englewood Cliffs, New Jersey: Prentice Hall.

[Salton 1970]

G. Salton. (1970). Automatic Processing of Foreign Language Documents. *Journal of the American Society for Information Science*, 21(3), 187-194.

[Salton 1971a]

G. Salton. (1971a). Relevance Feedback and the Optimisation of Retrieval Effectiveness. In G. Salton (Ed.), *The SMART Retrieval System. Experiments in Automatic Document Processing* (pp. 324-336). New Jersey: Prentice Hall.

[Salton 1971b]

G. Salton. (1971b). *The SMART Retrieval System - Experiments in Automatic Document Processing*. Englewood Cliffs, New Jersey: Prentice Hall.

[Salton 1997]

G. Salton & C. Buckley. (1997). Improving Retrieval Performance by Relevance Feedback. In K. Sparck Jones & P. Willet (Eds.), *Readings in Information Retrieval* (pp. 355-364). San Francisco, California: Morgan Kauffmann.

[Salton 1985]

G. Salton, E. Fox & E. Voorhees. (1985). Advanced Feedback Methods in Information Retrieval. *Journal of the American Society for Information Science*, 36(3), 200-210.

[Salton 1983]

G. Salton, E. Fox & H. Wu. (1983). Extended Boolean Information Retrieval. *Communications of the ACM*, 26(11), 1022-1036.

[Sanderson 1996]

M. Sanderson. (1996). *Word Sense Disambiguation and Information retrieval*. PhD Thesis, University of Glasgow.

[Savoy 2003]

J. Savoy. (2003). Report on CLEF2002 Experiments: Combining Multiple Sources of Evidence. In C. Peters (Ed.), *Advances in Cross-Language Information Retrieval - Third Workshop of the Cross-Language Evaluation Forum, CLEF 2002 (LNCS 2785)* (pp. 31-45): Springer.

[Schamber 1994]

L. Schamber. (1994). Relevance and Information Behavior. *Annual Review of Information Science and Technology*, 29, 3-48.

[Sheridan 1996]

P. Sheridan & J. P. Ballerini. (1996). Experiments in Multilingual Information Retrieval using the SPIDER System. In *19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. Zurich, Switzerland: ACM.

[Sheridan 1997]

P. Sheridan, M. Braschler & P. Schäuble. (1997). Cross-Language Information Retrieval in a Multilingual Legal Domain. In C. Peters & C. Thanos (Eds.), *Research and Advanced Technology for Digital Libraries: First European Conference, ECDL'97* (pp. 253-268). Pisa, Italy.

[Snowball]

Snowball. Retrieved on 29/07/2003, from
<http://snowball.tartarus.org/portuguese/voc.txt>

[Spink 1994]

A. Spink. (1994). Term Relevance Feedback and Query Expansion: Relation to Design. In *17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. (pp. 81-90). Dublin, Ireland.

[Swanson 1971]

D. R. Swanson. (1971). Some Unexplained Aspects of the Cranfield tests of indexing performance factors. *Library Quarterly*, 41(223-228).

[SYSTRAN]

SYSTRAN. Retrieved on 05/10/2003, from <http://www.systransoft.com/>

[Taylor 1998]

K. Taylor & J. White (Eds.). (1998). *Predicting What MT is Good for: User Judgements and Task Performance*: Springer.

[Telcordia]

Telcordia. Retrieved on 06/03/2001, from <http://lsi.research.telcordia.com/>

[TREC]

TREC. *The Text Retrieval Conference*. Retrieved on 17/07/2003, from <http://trec.nist.gov>

[University of Tennessee]

University of Tennessee. Retrieved on 18/07/03, from <http://www.cs.utk.edu/~lsi/>

[van Rijsbergen 1979]

C. J. van Rijsbergen. (1979). *Information Retrieval* (2nd ed.). London, UK: Butterworths.

[Voorhees 1998]

E. Voorhees. (1998). Variations in Relevance Judgements and the Measurement of Retrieval Effectiveness. In W. B. Croft (Ed.), *Proceedings of the 21st Annual International ACM-SIGIR Conference on Research and Development in Information Retrieval* (pp. 315-323): ACM Press.

[Voorhees 1997]

E. Voorhees & D. Harman. (1997). Overview of the Sixth Text Retrieval Conference. In E. Voorhees & D. Harman (Eds.), *The Sixth Text Retrieval Conference (TREC-6)*. Gaithersburg, Maryland: NIST.

[Voorhees 2001]

E. Voorhees & D. Harman. (2001). Overview of TREC 2001. In E. Voorhees & D. Harman (Eds.), *The Ninth Text Retrieval Conference (TREC-9)* (pp. 1-14). Gaithersburg, Maryland: NIST.

[Wang 2001]

J. Wang & D. W. Oard. (2001). iCLEF 2001 at Maryland. In C. Peters, M. Braschler, J. Gonzalo & M. Kluck (Eds.), *Evaluation of Cross-Language Information Retrieval Systems. Second Workshop of the Cross-Language Evaluation Forum, CLEF 2001. Lecture Notes in Computing Science 2406*: Springer.

[Wechsler 1997]

M. Wechsler, P. Sheridan & P. Schäuble (Eds.). (1997). *Multi-Language Text Indexing for Internet Retrieval*. Montreal, Canada.

[Williamson 1978]

R. Williamson. (1978). Does Relevance Feedback Improve Documents Retrieval Performance? In *1st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*: ACM.

[Witten 1999]

I. Witten, A. Moffat & T. C. Bell. (1999). *Managing Gigabytes - Compressing and Indexing Documents and Images* (2nd ed.). San Francisco, California: Morgan Kauffmann.

[Xu 1996]

J. Xu & W. B. Croft. (1996). Query Expansion Using Local and Global Document Analysis. In *19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 4-11).

[Yang 1997]

Y. Yang, J. Carbonell, R. D. Brown & R. Frederking. (1997). *Translingual Information Retrieval*. Paper presented at the 15th International Joint Conference on Artificial Intelligence (IJCAI), Nagoya, Japan.

[Zakis 1997]

J. D. Zakis & Z. J. Pudlowski. (1997). The World Wide Web as Universal Medium for Scholarly Publication, Information Retrieval and Interchange. *Global Journal of Engineering Education*, 1(3).

[Zelikowitz 2001]

S. Zelikowitz & H. Hirsh. (2001). Using LSI for Text Classification in the Presence of Background Text. In *Proceedings of the 10th International Conference on Information and Knowledge Management*.

[Zobel 1998]

J. Zobel. (1998). How Reliable are the Results of Large-Scale Information Retrieval Experiments? In W. B. Croft, A. Moffat, C. J. van Rijsbergen, R. Wilkinson & J. Zobel (Eds.), *21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 307-314). Melbourne, Australia: ACM.

APPENDIX A – PORTUGUESE STEMMING RULES

This appendix contains the stemming rules for Portuguese defined by the RSLP stemming algorithm described in Chapter 3. Each rule consists of a suffix to remove, a minimum size for the stem, a replacement suffix (when applicable), and a set of exceptions (when applicable).

Plural Reduction Rules

Suffix to Remove	Min Stem Size	Replacement	Exceptions	Example
ns	1	m		bons → bom
ões	3	ão		balões → balão
ães	1	ão	mãe	capitães → capitão
ais	1	al	cais, mais	normais → normal
éis	2	el		papéis → papel
eis	2	el		amáveis → amável
óis	2	ol		lençóis → lençol
is	2	il	lápiz, crúcis, biquínis, pois, depois, dois	barris → barril
les	3	l		males → mal
res	3	r		mares → mar
s	2		aliás, pires, mas, menos, férias, fezes, pêsames, gás, atrás, moisés, através, convés, ês, país, após, ambas, ambos, messias	casas → casa

Feminine Reduction Rules

Suffix to Remove	Min Stem Size	Replacement	Exceptions	Example
ona	3	ão	abandona, lona, iona, cortisona, monótona, maratona, acetona, detona, carona	chefona → chefão
ã	2	ão	amanhã, arapuã, fã, divã	vilã → vilão
ora	3	or		professora → professor
na	4	no	guiana, campana, grana, caravana, banana, paisana	americana → americano
inha	3	inho	rainha, linha, minha	sozinha → sozinho
esa	3	s		inglesa → inglês

Suffix to Remove	Min Stem Size	Replacement	Exceptions	Example
osa	3	oso	mucosa, prosa	famosa → famoso
íaca	3	íaco		maníaca → maníaco
ica	3	ico	dica	prática → prático
ada	2	ado	pitada	mantida → mantido
ida	3	ido	vida	caida → caído
ída	3	ido	recaída, saída, dúvida	cansada → cansado
ima	3	imo	vítima	prima → primo
iva	3	ivo	beira, cadeira, frigideira, bandeira, feira, capoeira, barreira, fronteira, besteira, poeira	passiva → passivo
eira	3	eiro	beira, cadeira, frigideira, bandeira, feira, capoeira, barreira, fronteira, besteira, poeira	primeira → primeiro

Adverb Reduction Rules

Suffix to Remove	Min Stem Size	Replacement	Exceptions	Example
mente	4		experimente	felizmente → feliz

Augmentative/Diminutive Reduction Rules

Suffix to Remove	Minimum Stem Size	Replacement	Exceptions	Example
díssimo	5			cansadíssimo → cansad
abilíssimo	5			amabilíssimo → ama
íssimo	3			fortíssimo → fort
ésimo	3			
érrimo	4			chiquérrimo → chiqu
zinho	2			pezinho → pe
quinho	4	c		maluquinho → maluc
uinho	4			amiguinho → amig
adinho	3			cansadinho → cansad
inho	3		caminho, cominho, carinho	carrinho → carr
alhão	4			grandalhão → grand
uça	4			dentuça → dent

Suffix to Remove	Min Stem Size	Replacement	Exceptions	Example
ação	4		antebraço	ricaço → ric
adão	4			casadão → cans
ázio	3		topázio	corpázio → corp
arraz	4			pratarraz → prat
arra	3			bocarra → boc
zão	2		coalizão	calorzão → calor
ão	3		camarão, chimarrão, canção, coração, embrião, grotão, glutão, ficção, fogão, feição, furacão, gamão, lampião, leão, macacão, nação, órfão, órgão, patrão, portão, quinhão, rincão, tração, falcão, espião, mamão, folião, cordão, aptidão, campeão, colchão, limão, leilão, melão, barão, milhão, bilhão, fusão, cristão, ilusão, capitão, estação, senão	meninão → menin

Noun Reduction Rules

Suffix to Remove	Min Size	Replacement	Example
encialista	4		existencialista → exist
alista	5		minimalista → minim
agem	3	coragem, chantagem, vantagem, carruagem	contagem → cont
iamento	4		gerenciamento → gerenc
amento	3	firmamento, fundamento, departamento	monitoramento → monit
imento	3		nascimento → nasc
alizado	4		comercializado → comerci
tizado	4	alfabetizado	traumatizado → trauma
izado	5	organizado, pulverizado	alfabetizado → alfabet
ativo	4	pejorativo, relativo	associativo → associ
tivo	4	relativo	contraceptivo → contracep
ivo	4		esportivo → esport
ado	2	grado	abalado → abal
ido	3	cândido, consolidado, rápido, decido, tímido, duvido, marido	impedido → imped

Suffix to Remove	Min Size	Replacement	Exceptions	Example
ador	3			ralador → ral
edor	3			entendedor → entend
idor	4			cumpridor → cumpr
atória	5			obrigatória → obrig
tor	2		benfeitor, leitor, editor, pastor, produtor, promotor, consultor	produtor → produt
abilidade	5			comparabilidade → compar
icionista	4			abolicionista → abol
cionista	5			intervencionista → interven
ional	4			profissional → profiss
ência	3			referência → refer
ância	4		ambulância	repugnância → repugn
edouro	3			abatedouro → abat
queiro	3	c		fofoqueiro → fofoc
eiro	3			brasileiro → brasil
oso	3	precioso		gostoso → gost
alizaç	5			comercializaç → comerci
ismo	3			consumismo → consum
izaç	5		organizaç	concretizaç → concret
aç	3		equaç, relaç	alegaç → aleg
iç	3		eleição	aboliç → abol
ário	3		voluntário, salário, aniversário, diário, lionário, armário	anedotário → anedot
ério	6			ministério → minist
ês	4			chinês → chin
eza	3			beleza → bel
ez	4			rigidez → rigid
esco	4			parentesco → parent
ante	2		gigante, elefante, adiante, possante, instante, restaurante	ocupante → ocup
ástico	4		eclesiástico	bombástico → bomb
ático	3			problemático → problem
ico	4		político, eclesiástico, diagnostico, prático, doméstico, diagnóstico, idêntico, alopático, artístico, autêntico, eclético, crítico, critico, público, explico	polêmico → polêm
ividade	5			produtividade → produt
idade	5		autoridade, comunidade	profundidade → profund
oria	4		categoria	aposentadoria → aposentad
encial	5			existencial → exist

Suffix to Remove	Min Size	Replacement	Exceptions	Example
ista	4			artista→art
quice	4	c		maluquice→maluc
ice	4		cúmplice	chatice→ chat
íaco	3			demoníaco→demon
ente	4		frequente, alimento, crescente, permanente, oriente, aparente	decorrente→decorr
inal	3			criminal→crim
ano	4			americano→ americ
ável	2		afável, razoável, potável, vulnerável	amável→ am
ível	5		possível	combustível→combust
ura	4		imatura, acupuntura, costura	cobertura→cobert
ual	3			consensual→consens
ial	3			mundial→mund
al	4		afinal, animal, estatal, bissexual, desleal, fiscal, formal, pessoal, liberal, postal, virtual, visual, pontual, sideral, sucursal	experimental→experiment
ismo	4		cinismo	nativismo→nativ

Verb Reduction Rules

Suffix to Remove	Min Size	Replacement	Exceptions	Example
aríamo	2			cantaríamo → cant
ássemo	2			cantássemo → cant
eríamo	2			beberíamo → beb
êssemo	2			bebêssemo → beb
iríamo	3			partiríamo → part
íssemo	3			partíssemo →part
áramo	2			cantáramo→ cant
árei	2			cantárei→ cant
aremo	2			cantaremo → cant
ariam	2			cantariam→ cant
aríei	2			cantaríei→ cant
ássei	2			cantássei→ cant
assem	2			cantassem→ cant
ávamo	2			cantávamo→ cant
êramo	3			bebêramo→ beb
eremo	3			beberemo→ beb
eriam	3			beberiam→ beb
eríei	3			beberíei→ beb
êssei	3			bebêssei→ beb
essem	3			bebessem→ beb

Suffix to Remove	Min Size	Replacement	Exceptions	Example
íramo	3			partiríamo→ part
iremo	3			partiremo→ part
iriam	3			partiriam→ part
iríei	3			partiríei→ part
íssei	3			partíssei→ part
issem	3			partissem→ part
ando	2			cantando→ cant
endo	3			bebendo→ beb
indo	3			partindo→ part
ondo	3			propondo→ prop
aram	2			cantaram→ cant
arde	2			cantarde → cant
arei	2			cantarei→ cant
arem	2			cantarem→ cant
aria	2			cantaria→ cant
armo	2			cantarmo→ cant
asse	2			cantasse → cant
aste	2			cantaste→ cant
avam	2		agravam	cantavam→ cant
ávei	2			cantávei→ cant
eram	3			beberam→ beb
erde	3			beberde→ beb
erei	3			beberei→ beb
êrei	3			bebêrei→ beb
erem	3			beberem→ beb
eria	3			beberia→ beb
ermo	3			bebermo→ beb
esse	3			bebesse→ beb
este	3		faroeste, agreste	bebeste→ beb
íamo	3			bebíamo→ beb
iram	3			partiram→ part
íram	3			concluíram→ conclu
irde	2			partirde→ part
irei	3		admirei	partirei→ part
irem	3		adquirem, admirem	partirem→ part
izar	5		organizar,	concretizar→concret
itar	5		acreditar, explicitar, estreitar	debilitar→debil
iria	3			partiria→ part
irmo	3			partirmo→ part
isse	3			partisse→ part
iste	4			partiste→ part
amo	2			cantamo→ cant
ara	2		arara, prepara	cantara→ cant
ará	2		alvará	cantará→ cant
are	2			cantare→ cant
ava	2		agrava, iava	cantava→ cant

Suffix to Remove	Min Size	Replacement	Exceptions	Example
emo	2			cantemo→cant
era	3			bebera→beb
erá	3		acelera, espera	beberá→beb
ere	3		espere	bebere→beb
iam	3		enfiam, ampliam, elogiam, ensaiam	bebiam→beb
íei	3			bebíei→beb
imo	3		reprimio, intimo, íntimo, nimo, queimo, ximo	partimo→part
ira	3		fronteira, sátira	partira→part
irá	3			partirá→part
ire	3		adquire, admire	partire→part
omo	3			compomo→comp
ai	2			cantai→cant
am	2			cantam→cant
ear	4		alardear, nuclear	barbear→barb
ar	2		azar, bazaar, patamar	cantar→cant
uei	3			cheguei → cheg
ei	3			cantei→cant
guem	3	g		alaguem→alag
em	2		alem, virgem	cantem→cant
er	2		éter, pier	beber→beb
eu	3		chapeu	bebeu→beb
ia	3		estória, fatia, acia, praia, elogio, mania, lábia, aprecia, polícia, arredia, cheia, ásia	bebia→beb
ir	3		freir	partir→part
iu	3			partiu→part
ou,	3			chegou→cheg
i	3			Bebi→beb

Thematic Vowel Removal

Suffix to Remove	Min Stem Size	Replacement	Exceptions	Example
a	3			menina →menin
e	3			grande →grand
o	3			menino →menin

APPENDIX B – RELEVANCE FEEDBACK QUESTIONNAIRE

This Appendix contains the questionnaire that was answered by the participants in the Relevance Feedback experiment described in Chapter 5. The first two pages detail the instructions as originally given to the subjects.

Início

Fim

Introdução

Esta pesquisa é parte de uma tese de doutorado que está sendo realizada na Middlesex University (Inglaterra). O objetivo é avaliar o desempenho de sistemas de recuperação de informações e não o desempenho das pessoas que respondem a pesquisa.

Obrigado por sua participação.

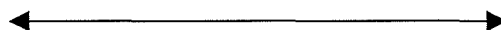
Instruções

- A tarefa consiste em ler a descrição de um tópico de consulta. Imagine que você está fazendo uma pesquisa e irá escrever um relatório sobre o tópico. Então consulta um site de busca que lhe retorna uma série de documentos. A sua tarefa é então julgar cada documento com relação ao tópico como “relevante”, “irrelevante” ou “não sei”.
- Documentos em que somente alguma parte é relevante ao tópico devem ser julgados como relevante.
- Os documentos devem ser julgados individualmente, isto é, a relevância ou irrelevância de um documento não deve afetar os demais.
- A tarefa é composta por 6 consultas. Para cada consulta você irá ler 10 documentos.
- Os documentos são notícias publicadas pelo jornal Los Angeles Times no ano de 1994.
- Não há limite de tempo.
- Os julgamentos devem ser feitos individualmente e sem o uso de dicionários.
- Você não precisa ler os textos na íntegra. Leia o necessário para fazer os julgamentos.

Antes de prosseguir por favor responda:

Em uma escala de 1 a 5, como você classifica a sua habilidade de escrever em inglês (onde 1 significa “fluente” e 5 significa “nehuma”):

Fluente 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Nenhuma



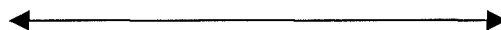
Como você classifica a sua habilidade de ler em inglês (onde 1 significa “fluente” e 5 significa “nenhuma”):

Fluente 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Nenhuma



Com que frequência você costuma utilizar mecanismos de busca (Google, Altavista, Yahoo, etc.) na Internet (onde 1 significa "semanalmente" e 5 significa "raramente"):

Semanalmente 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Raramente



Sanções da ONU contra o Iraque

Que medidas tem tomado o Iraque para a suspensão do embargo econômico da ONU e as sanções políticas impostas depois da sua invasão do Kuwait em 1990?

Documentos devem incluir formas pelas quais o Iraque tenha tentado a suspensão das sanções. Simples descrições das sanções ou retórica contra as sanções não são relevantes. Expressões de arrependimento do Iraque por ter invadido o Kuwait são relevantes

1.

(91840)<DOCNO> LA100894-0025

IRAQ BACK TO SQUARE 1 ON SANCTIONS; MIDEAST: BUILDUP OF TROOPS COUNTERPRODUCTIVE TO EFFORTS TO GET U.N. TO LIFT CURBS THAT HAVE CRIPPLED THE NATION'S ECONOMY.

By sending some of his best troops south toward the border with Kuwait, Iraqi President Saddam Hussein seems to have shattered his country's hope of having the United Nations lift the sanctions that have devastated its economy.

Word of the troop movements came Friday as Iraq's top diplomat, Deputy Prime Minister Tarik Aziz, was appealing to the United Nations to lift the sanctions, which he said are responsible for growing malnutrition, inadequate medical care and widespread suffering.

In a speech to the U.N. General Assembly devoted almost entirely to the issue of sanctions, Aziz claimed that Baghdad has fully implemented many of the conditions imposed by the U.N. Security Council after Iraq's defeat in the 1991 Persian Gulf War and is entitled to at least partial relief from the sanctions.

But he said the United States has prevented the council from relaxing the embargo.

From Iraq's standpoint, the most troublesome sanction prohibits Baghdad from selling its oil on world markets except under a tight U.N. supervision in which all of the revenue would be devoted to humanitarian relief within Iraq and to payment of war reparations to Kuwait. Iraq denounced the restrictions as a violation of its sovereignty and refused to sell its oil under those conditions.

Iraq offered no explanation for advancing its forces toward Kuwait. A Defense Ministry statement said only that Iraqi troop movements on Iraqi territory were no one else's business. But coupled with Aziz's speech in New York, the action seemed a particularly heavy-handed -- and almost certainly self-defeating -- attempt to intimidate the world community.

"If it is a serious deployment, then it runs counter to any interest in lifting the embargo," said William B. Quandt, a former National Security Council Middle East expert.

"If they are trying to get countries to ease up on the embargo, it seems counterproductive, because one requirement for lifting the embargo is for them to respect the Kuwait border," said Quandt, now a professor at the University of Virginia.

The sanctions originally were imposed after Iraq's invasion of Kuwait in August, 1990. After Iraq's defeat the following year in the Persian Gulf War, the Security Council ruled that the sanctions would remain in place until Iraq destroyed its nuclear and chemical weapons programs, ended persecution of Kurds and other Iraqi minority groups, recognized the sovereignty of Kuwait and accepted a U.N.-imposed settlement of a border dispute with Kuwait.

Before the latest troop movements, Iraq seemed to be making some headway in its effort to win the right to unrestricted sale of its oil. Earlier this year, U.N. inspectors reported that Iraq was in substantial compliance with the provisions requiring destruction of its nuclear weapons potential.

French Foreign Minister Alain Juppe said recently that his country -- like the United States a permanent member of the Security Council -- believes that it is time to allow Iraq to resume oil sales, at least for a trial period.

But the United States continues to insist that the sanctions be kept in place, saying the Iraqis have not complied with all aspects of the resolution. Iraq has not yet officially recognized Kuwait or accepted the border demarcation.

In his speech to the General Assembly, Aziz bitterly attacked the U.S. position.

"The United States of America is persistently obstructing any steps toward the direct legal application of the resolutions of the (Security) Council, particularly those provisions relating to lifting sanctions from Iraq, and is conducting itself on the basis of biased political motives which bear no relation to either the resolutions of the Council or the (U.N.) charter," he said.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

2.

(7077)<DOCNO> LA042894-0255

U.S. SEEKS TO HOLD LINE ON IRAQ EMBARGO; MIDEAST: CHRISTOPHER, GULF STATES PRESENT UNITED FRONT AGAINST PRESSURE TO EASE SANCTIONS.

The United States and six Persian Gulf monarchies launched a counteroffensive Wednesday against growing international pressure to ease economic sanctions against Iraq, just one day after Baghdad agreed to accelerate steps required to lift the three-year embargo against it.

On the first stop of his Middle East swing, Secretary of State Warren Christopher held talks with six foreign ministers of the Gulf Cooperation Council to "unanimously and strongly express our common resolve to maintain the sanctions until Iraq complies fully with all U.N. resolutions." But three of five permanent members of the U.N. Security Council -- France, Russia and China -- are either pushing for a schedule to lift the oil embargo as the first step in ending the broader sanctions or are neutral on the issue.

France and Russia are reportedly interested in exploring lucrative trade deals with Baghdad after the oil embargo is removed.

Turkey, which borders Iraq and was a key member of the U.S.-led coalition in the 1991 Persian Gulf War, has also been pressuring the United States to ease the embargo. Use of Turkey's oil pipeline from Iraq to the Mediterranean has been frozen for three years, costing more than \$100 million and further crippling Ankara's troubled economy.

The embargo, which is reviewed every 60 days, is due for discussion at the United Nations again early next month. U.S. officials have been concerned that the push to end the oil embargo could lead to a phased plan that would ease the political and economic squeeze on the regime of President Saddam Hussein. "Once you start talking about deadlines or ways to end the embargo, then the game is up," one senior U.S. official said.

Among many ongoing violations, Christopher specifically cited Iraq's suppression of its own citizens; its refusal to recognize Kuwait's independence or borders, and its promotion of terrorism, most recently in the assassination this month of a leading Iraqi dissident in Beirut.

"Our resolve to defend this interest against aggression from any quarter is as strong today as it was when the United States led the coalition effort to liberate Kuwait from Iraq's aggression," Christopher said at a news conference here. "We remain vigilant and determined. We are conscious that the region is far from free of threats." The talks in Riyadh came a day before a report is due at the Security Council from Rolf Ekeus, chairman of the U.N. Special Commission on Iraq, on Baghdad's progress in complying with various U.N. resolutions.

After talks in Iraq on Tuesday, Ekeus and Iraqi Deputy Prime Minister Tarik Aziz agreed to expedite efforts to create an arms-monitoring program to prevent Baghdad from developing weapons of mass destruction, a key condition for lifting sanctions.

A joint statement expressed appreciation from the Special Commission and the International Atomic Energy Agency "for Iraq's expeditious and positive responses to their requests for assistance and support." Ekeus has also said recently that Iraq is cooperating with U.N. monitors, in contrast to its obstruction of the inspectors in the early days of their work.

In recent weeks, the United Nations has increased the number of monitors inspecting Iraqi installations and has begun installing equipment, including cameras and sensors, at critical facilities.

But U.S. officials say they are very skeptical about Iraq's compliance and intentions. Baghdad "retains the desire to threaten its neighbors and U.S.

interests (and) has no apparent intention of actually giving up its determination to retain its weapons of mass destruction," Ronald Neumann, the State Department's director of Gulf affairs, said recently in a speech in Washington.

U.S. officials believe the next 18 to 24 months will be critical in determining if the Hussein regime can survive economically -- and therefore politically. At the time of the 1990 invasion of Kuwait, Iraq was estimated to have between \$5 billion and \$7 billion in foreign exchange reserves. But because of the sanctions, Baghdad is now believed to have gone through much of its reserves. Its oil-based economy is deteriorating seriously, with widespread shortages of basic goods.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

3.

(106613)<DOCNO> LA120594-0182

BAGHDAD WRONG TO INVADE KUWAIT, IRAQI OFFICIAL SAYS

A top Iraqi official has told a Swedish newspaper that the 1990 invasion of Kuwait was a serious mistake.

The statement by Barzan Ibrahim Tikriti, a half brother of Saddam Hussein and head of Iraq's U.N. mission in Geneva, came in an exclusive interview published Sunday by Svenska Dagbladet. It was the first time a high-level official has admitted publicly that the invasion was a mistake.

"It's regrettable that Iraq invaded Kuwait," Tikriti said. "I wish it would not have happened since the consequences have been terrible for the Iraqi people." He expressed frustration over the United States' unwillingness to talk to Iraq, which last month officially recognized Kuwait's borders and independence.

"They refuse to see us, but the United States is willing to meet with dictators like North Korea's," he said.

"We do everything we can to put things right again. The United States and Iraq cannot continue to be enemies; we need each other." Tikriti is the youngest of Hussein's three half brothers from his mother's second marriage. His father is a brother of Hussein's father.

Tikriti has had his ups and downs with Hussein, but the blood relationship seems to have endured. According to Western intelligence agencies and Iraqi defectors, he heads Hussein's intelligence network in Europe, one of the key links in the Iraqi leader's terror apparatus, and handles the family's investments abroad.

"It is time for the United States and Iraq to normalize relations," he said in the interview. "We want to have an open and honest dialogue with the Americans.

. . . There is no longer need for tough sanctions." The United States has led opposition to the lifting of U.N. sanctions imposed on Iraq following its invasion of Kuwait in August, 1990.

Tikriti has been shunned by the diplomatic community in Geneva, which considers him little more than a thug dressed as a diplomat. He once beat up his chauffeur at a diplomatic reception when the man was late picking Tikriti up.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

4.

(94689)<DOCNO> LA101994-0006

IRAQI COMPLIANCE, NOT SWEET TALK; COPING WITH HUSSEIN'S HEGEMONIC AMBITIONS

Late last Saturday, after a dozen hours of pointed argument and negotiation, the U.N. Security Council unanimously approved a resolution condemning Iraq's movement of troops close to the Kuwait border and demanding not only their withdrawal but also that Iraq never again threaten its neighbors. Less than 48 hours later the consensus that put this resolution on record had fractured.

Russia, which had made no effort to mask its dissatisfaction with the resolution it found expedient to support Saturday night, discovers itself in deepening disagreement with the United States and to a lesser extent Britain over where U.N. policy toward Iraq should go next.

Russia's view, as expressed by Foreign Minister Andrei V. Kozyrev, is that the United States overreacted to what Baghdad intended to be innocent and routine military maneuvers in southern Iraq. Moreover, Iraq has made every effort to open up certain of its most suspect military facilities to U.N. inspections.

Further, it says it is conditionally prepared to recognize Kuwait's independence, dropping its claim that Kuwait is a separated province of Iraq.

Given this moderation, Moscow says, Iraq deserves to be rewarded. Russia wants the economic sanctions against Iraq eased. Specifically, it wants the ban on its oil sales lifted.

The American response, as forcefully expressed by Ambassador Madeleine Albright, is that "words are cheap. Actions are the coin of the realm." Saddam Hussein trails a long and sordid record of meaningless promises and bogus assurances that can't be erased by yet more consoling lies and deceptive sweet talk. What Washington demands is full Iraqi compliance -- actions, not words -- with relevant U.N. resolutions. For starters, that means unequivocal recognition of Kuwait's independence and international borders. That, along with other appropriate unconditional actions, must precede any end to the embargo.

Iraq's rubber-stamp parliament had the chance a few days ago to renounce the claim to Kuwait. It didn't do so. Its deputy prime minister, Tarik Aziz, had the opportunity to underscore Iraq's honorable intentions when he spoke to the Security Council on Monday. He did not do so.

Russia has powerful commercial reasons -- Baghdad has promised it both debt repayment and lucrative contracts -- for seeking a softer approach toward Iraq.

But Washington has a more compelling motivation for insisting that Iraq abide by the U.N. resolutions: Neither this country nor its allies want to fight a second Persian Gulf war in response to Saddam Hussein's hegemonic ambitions.

That really is what the debate over lifting the sanctions is about.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

5.

(100456)<DOCNO> LA111194-0058

IRAQ OFFICIALLY RECOGNIZES KUWAIT, BUT U.S. IS SKEPTICAL; MIDEAST: HUSSEIN USED CLAIM TO LAND AS EXCUSE FOR 1990 INVASION . WASHINGTON NOT READY TO LIFT SANCTIONS.

Iraq, heeding the demand of the Security Council, officially recognized Kuwait on Thursday, but skeptical American officials insisted that Iraqi leader Saddam Hussein must still meet a host of other requirements before the United Nations can lift any sanctions.

The recognition means that Iraq has abandoned its traditional claim to Kuwait, a claim that Hussein invoked as his excuse for invading the small emirate in 1990 and precipitating the Persian Gulf War that ended with his defeat.

According to a statement issued in Baghdad, the Iraqi capital, and purportedly signed by Hussein, the Revolutionary Command Council decided to recognize "the sovereignty of the state of Kuwait, its territorial integrity and political independence." The council, headed by Hussein, also recognized the international Iraq-Kuwait boundary set down by a U.N. commission in 1991.

The council said it was taking these steps because of its respect for the U.N.

Charter and international law and its intention "to comply fully with all relevant resolutions of the United Nations Security Council." The Iraqi National Assembly also recognized the independence and boundary of Kuwait, passing the resolution in a special session attended by Russian Foreign Minister Andrei V. Kozyrev. The Russian foreign minister has annoyed American officials recently with his persistent campaign to lift some of the sanctions on Iraq.

The Iraqi News Agency quoted Kozyrev as saying, "There is a wall of ice surrounding Iraq, intended to isolate it, but the sun has now risen to melt this ice." Despite the skeptical reaction from Americans, the Iraqi government intends to press its case for at least a partial end to sanctions by sending Deputy Prime Minister Tarik Aziz, who has long served Hussein as spokesman to the Western world, to the United Nations on Monday. American Ambassador Madeleine Albright told reporters that she has agreed to meet him, but only in her capacity as president of the Security Council for the month of November.

A meeting of the two would mark the highest-level Iraqi-American meeting since the Gulf War. But a U.S. official cautioned against "attaching any significance to the meeting." Albright will not be meeting Aziz as American ambassador, he said.

James P. Rubin, spokesman for Albright, made it clear that the United States will not agree to any lifting of sanctions when the Security Council makes its bimonthly review of the issue next week. Without the acquiescence of the United States, which has the power of veto in the Security Council, no lifting of sanctions is possible.

The recognition of Kuwait "complied with only one of a number of requirements set by the Security Council," Rubin said. "We have to keep in mind that Saddam Hussein's blatant provocation against Kuwait and the United Nations last month demonstrated anew his untrustworthiness and his readiness to resort to threats and confrontation to achieve his objectives.

"It will be important," he went on, "that the council continue to make clear there will be no modification of the sanctions regime so long as Iraq has failed to establish its peaceful intentions by complying fully with all relevant Security Council resolutions." This longstanding American position is at odds with that of Kozyrev and, to some extent, the French government. The Russians and French have argued that the Security Council should follow the strict letter of the Security Council resolutions. Under this strict reading, the sanctions on the sale of oil will be lifted when a U.N. special commission determines that Iraq has eliminated all of its weapons of mass destruction and has submitted to a system of long-term monitoring that makes sure it does not build such weapons again.

But American officials say that it is impossible to trust Hussein's acceptance of a monitoring system as long as he has not complied with all Security Council resolutions. To prove his good intentions, they say, the Iraqi leader must comply with various other Security Council demands, including an accounting of missing Kuwaiti prisoners, a halt in the persecution of the Kurds and Shiite Muslims in his country, and payment of compensation to Kuwait for destroyed property and to other countries hurt by the war.

The argument has led to some angry exchanges between the Americans, who appear sure that the French and Russians are interested only in resuming a lucrative trade, and the French and Russians, who seem certain that the Americans have no intention of ever lifting sanctions against Iraq.

The American refusal to halt sanctions was strengthened by a report from Rolf Ekeus of Sweden, the head of the special commission, who said Iraqi officials barely cooperated with his inspectors during the past month.

But Ekeus attributed this to fallout from the crisis provoked by Hussein last month when he moved troops to the Kuwait border in a manner that the United States said threatened a new invasion.

Ekeus told reporters that he looked on the recognition of Kuwait as an "extremely positive" development because it seemed logical for it to lead to "a policy of full cooperation" with the commission's inspectors.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

6.

(56920)<DOCNO> LA051994-0012

STAY ON THE MORAL HIGH GROUND; U.S. IS RIGHT THAT IRAQ MUST DO MORE IF SANCTIONS ARE TO BE LIFTED

The economic sanctions imposed on Iraq after its invasion and brutal occupation of Kuwait are still in force, but support for the measures is faltering. Three permanent members of the U.N. Security Council -- China, France and Russia -- have signaled they are ready to think about easing the restrictions, which are now in their fourth year. All, surely by no coincidence, have reportedly lined up contracts to help rebuild Iraq's infrastructure once commerce with Baghdad again is made legal.

The United States and Britain, the other two permanent members of the council, rightly insist that Iraq must comply more fully with U.N. resolutions before any changes are made.

Specifically they say Saddam Hussein's regime must recognize Kuwait's independence and stop waging ruthless war against Iraqi Kurds and Shiites.

These are valid and honorable demands, and Washington shouldn't hesitate to use its Security Council veto to enforce them.

Iraq does appear to have complied with U.N. orders to halt and destroy its programs to build nuclear and chemical weapons; whether it has also done so with its biological weapons program is less clear. Iraq argues that by ending its production of unconventional weapons it has earned the right to be freed from sanctions. But Washington and others question whether Iraq has in fact made the decision to adopt more responsible policies, domestically and in its foreign relations.

Iraq could hardly have resisted the internationally supervised destruction of its chemical and nuclear weapons programs without inviting harsh reprisals, since defiance of the Security Council would have constituted an unmistakable threat to peace and security.

However, many U.N. members appear ready to regard Iraq's actions against its Kurdish and Shiite citizens and its refusal to accept Kuwait's independence as matters of much less international concern. The United States and Britain, taking the moral high ground, don't see it that way at all. They are right to continue demanding that Iraq must end its internal persecution and recognize Kuwait's border before the sanctions are lifted.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

7.

(93397)<DOCNO> LA101394-0365

DEPLOYMENT HAS BACKFIRED, IRAQ ENVOY SAYS

Iraq's decision to deploy thousands of troops near its border with Kuwait backfired and has dealt "a big -- but not insurmountable -- setback" to Baghdad's effort to get out from under the United Nations' crushing economic embargo, Nizar Hamdoun, Iraqi ambassador to the United Nations, said Wednesday.

Hamdoun said that President Saddam Hussein acted without informing his top foreign policy adviser, Deputy Prime Minister Tarik Aziz, who was in New York at the time trying to negotiate an end to the sanctions that have all but devastated Iraq's economy.

Hamdoun's remarks represented an extraordinarily candid assessment for a senior Iraqi official, as Hussein is a ruthless dictator who has repeatedly stamped out all dissent.

But U.S. experts said that the U.N. ambassador's statements, made in an interview Wednesday, reflect increasingly visible tension between hard-liners in Baghdad, who apparently urged Hussein to move the troops south, and the Foreign Ministry, which has been trying to end the embargo through negotiations.

Aziz and Hamdoun were at the U.N. ambassador's Manhattan apartment last Friday when they were shocked to see a report of the deployment on Cable News Network, Hamdoun said. The two had been working with Russian and French diplomats to try to shape an embargo proposal that would win U.N. support, and neither of them had any idea that the troops were on the move, he said.

Aziz was "livid" that the deployment had undermined their efforts at the United Nations, according to an American source who is in frequent contact with Iraqi officials.

The next day Aziz left for Baghdad, stopping briefly in Amman, Jordan, to confer with King Hussein. During the Persian Gulf War, the king sided with Baghdad, but in the current confrontation he has vowed to join in defending Kuwait.

In Baghdad, Aziz told Hussein that the United States was mounting a massive military buildup and was serious about confronting any Iraqi threat. The Iraqi leader, apparently persuaded, ordered the troops withdrawn.

State Department officials refused to comment on Hamdoun's account.

But a leading U.S. expert on Iraq, Phebe Marr of the National Defense University, said that increasingly visible disagreements have emerged between hard-liners in Baghdad and Foreign Ministry officials such as Hamdoun and Aziz.

"There is clearly a split between doves and hawks," she said. "The divide has been in the press in Baghdad for weeks." Although Hussein apparently remains in firm control, she said, "there is more debate going on, and more cracks in the regime than there used to be." Hamdoun and Aziz have been the most visible "doves," working for the last year to get the U.N. sanctions lifted and, now that "it has blown up in their faces, it's understandable that they might be upset," Marr said.

Most prominent among the "hawks," she said, is Hussein's son Uday, who publishes Babel, a daily newspaper.

"Babel has attacked Tarik Aziz directly," she said. "Uday is saying that the Foreign Ministry is wrong, that the negotiations can't work because the United States is not going to lift the sanctions and that the only correct strategy for Iraq is to raise the costs to the West of maintaining the sanctions." Aziz, a veteran foreign policy adviser, and Hamdoun, former Iraqi ambassador to the United States, both had extensive dealings with U.S. officials before Washington broke off relations with Baghdad in 1990 as a result of Iraq's invasion of Kuwait and the war.

Hamdoun, in an interview over lunch at a New York restaurant near his office, said Hussein's deployment of the troops was designed to get the attention of the U.N. Security Council, which has been monitoring Iraq's progress in meeting U.N. disarmament demands.

Iraq, he insisted, is in full compliance with U.N. disarmament demands dealing with nuclear, chemical and biological weapons. And until the deployment of the troops, he said, Russian and French diplomats had reacted positively to Iraq's efforts. But he pointed out that U.S. and British officials had adamantly opposed lifting the sanctions.

The United States and Britain maintain that Iraq has not fulfilled several of the conditions for lifting the embargo, including recognition of Kuwait's sovereignty and resolution of a dispute over the border. France and Russia acknowledge that, but argue that some relaxation of the embargo would encourage Hussein to meet the rest of the U.N. terms. They also argue that the suffering of Iraqi citizens, unable to affect their regime, is too great.

Hamdoun indicated that Iraq was also prepared to deal with the issues of Kuwait's border and its sovereignty rights and that he had received information that the Kuwaitis were interested in negotiating but that Washington had rejected the idea.

At one point as Hamdoun spoke, his cellular telephone rang. The Iraqi foreign minister, Mohammed Said Sahaf, was calling.

After a brief conversation, Hamdoun put down the phone. He looked relieved.

"He said all the Iraqi troops have been withdrawn from the area near Kuwait," Hamdoun said. "The Russian and Chinese military attaches are in the area seeing the troops removed and will report back to their governments." Hamdoun indicated that, with the massive U.S. military buildup, Baghdad is bracing for a punitive strike even though the Iraqi troops have withdrawn. He said his course will be to "just continue working with the French and Russians to try to get the sanctions lifted." Hamdoun served as Iraq's ambassador to Washington from 1983 to 1987, which he called the heyday of the two countries' relations. At that time, Washington was backing Iraq in its war with Iran.

Returning to Iraq, he served as deputy foreign minister until 1992 when he was named Iraq's ambassador to the United Nations. He now lives in New York with his wife and two daughters, 11 and 10 years old, both born in Washington.

Hamdoun expressed frustration over the fact that American officials refuse to deal with Iraqi officials even at the United Nations. He said that Aziz had called on U.S. Ambassador Madeleine Albright, but she had refused to see him.

Nelson reported from New York and McManus from Washington.

Em relação ao tópico, este documento é:

- ☐ relevante
☐ irrelevante
☐ não sei

8.

(16269)<DOCNO> LA092594-0179

IRAQ'S ARMS -- DON'T TRUST, DO VERIFY; U.N. SANCTIONS AGAINST OIL EXPORTS SHOULD BE LIFTED ONLY UNDER THE STRICTEST CONDITIONS

Iraq's biggest money-earner is oil, but ever since the 1991 Persian Gulf War that nation has been under U.N. Security Council-imposed sanctions that have kept its oil from being sold on world markets and prevented hard currency from flowing into Baghdad's treasury. Those sanctions would be lifted, the council said 3 1/2 years ago, once it was clear that Iraq no longer posed a security threat to its neighbors, especially Kuwait, which Iraq invaded and brutally occupied in 1990.

Now the council has been told by Rolf Ekeus, the diplomat it put in charge of monitoring Iraq's military status, that Saddam Hussein's regime has lost its capability to produce weapons of mass destruction -- nuclear, chemical or biological -- and that therefore the regional military threat it presents has been largely removed.

Assuming the oversight system put in place by the United Nations to make sure Iraq doesn't cheat on the weapons ban works, the council will probably be asked sometime next spring to allow Iraq once again to begin exporting oil. Under the right conditions -- and only under the right conditions -- the United States should support that request.

A key condition is that Baghdad give up its spurious claim that Kuwait is simply a lost province that Iraq has the right to reclaim by force. Kuwait is an internationally recognized sovereign country, as legitimate a political entity as Iraq itself, and Baghdad must recognize it as such, along with the redefined border between the two states that was fixed by the U.N. Boundary Commission in 1992.

A second condition is that Iraq agree to accept indefinite international monitoring of its imports, to prevent it from reacquiring the machinery and materials to rebuild its now liquidated unconventional weapons capability. An essential part of that oversight would be cooperation from all countries that do business with Iraq.

Saddam Hussein's regime came perilously close to setting up a nuclear weapons industry because it was able to obtain massive amounts of needed technology from outside, often by deceit, sometimes with the compliance of suppliers and export-hungry governments. That can't be allowed to happen again.

It of course remains a moral outrage that Hussein and his bloody-handed collaborators are still in power. But a change of government in Iraq was not a demand set by the Security Council for lifting the sanctions.

If Iraq is found no longer to be a threat to its neighbors, then the legal basis for the sanctions disappears. The foundation for that finding has now been laid, with a monitoring system to ensure Iraq doesn't rearm soon to be tested. If that system can be shown to work effectively, and with the additional conditions we noted above, the time will have come to lift the sanctions.

Em relação ao tópico, este documento é:

- ☐ relevante
☐ irrelevante
☐ não sei

9.

(92743)<DOCNO> LA101194-0223

HUSSEIN MAY BE BIDDING TO CURB INTERNAL WOES

Like a cornered cat, Iraqi President Saddam Hussein is making a desperate bid to survive growing internal pressures and prospects that the international sanctions strangling his economy will not be lifted soon, according to U.S.

experts.

In provoking a crisis -- massing his troops near the Kuwaiti border and then announcing a pullback -- he hopes to divert domestic attention from growing evidence of instability, marked by a spate of recent bombings of government targets, purges, the halving of rations and Draconian new punishments for crimes and defections.

"When he took over the premiership months ago, he promised things were going to get better. But instead the economy and the security situation only got worse," said Phebe Marr, an Iraq specialist at Washington's National Defense University.

"Then as it became apparent that the debate on U.N. sanctions was not going to yield much, Saddam couldn't be seen to be sitting in Baghdad doing nothing." The surprise move also marked an attempt to manipulate international opinion, perhaps in the naive belief that the United States might engage in negotiations to avoid a showdown while it is preoccupied with Haiti or as it has done with the regimes in North Korea and Cuba recently.

Yet the regime's pledge Monday to back down does not necessarily mark the end of the latest crisis -- nor the final phase of Hussein's rule. It reminds the international community that Hussein can still cause severe problems in the region -- and he knows it.

"The fact of the matter is that Iraq remains the military power in the region and is still a threat to its (Persian) Gulf neighbors," said Anthony H.

Cordesman, Georgetown University professor and author of the just published "Iran and Iraq: The Threat from the Northern Gulf." Baghdad could dispatch troops to the southern frontier whenever it wants -- precipitating a series of costly crises and playing the United States and its allies like a yo-yo.

"The problem isn't just whether he's dancing at the border but whether his forces are in southern Iraq. It takes days to get U.S. ships from (the Indian Ocean island of) Diego Garcia. There's always the risk of being overrun," Cordesman noted.

"He also may think he can pin down the United States indefinitely if we don't deal with him. He may think over time he can force us to negotiate to avoid a permanent or enduring confrontation." Hussein's unpredictability thus leaves the Clinton Administration with a thorny and potentially continuing problem. And so far, Saudi Arabia and other allies have not offered to pick up the tab for this operation, as they did to the tune of more than \$65 billion for Operations Desert Shield and Desert Storm.

But whatever the cost, the coalition probably would have to redeploy for other reasons unrelated to Kuwait's security, experts contend. Any instability in the region that would cut off the free flow of oil, for instance, could force the coalition into action.

"This is a wake-up call. Since the Gulf War, the United States has become more, not less, dependent on Gulf oil," said Geoffrey Kemp, senior associate at the Carnegie Endowment for International Peace.

"And there's nothing on the horizon to indicate any change over the next decade. So we and our allies have to pay more attention to the Gulf, not less." But in leading the anti-Hussein alliance, the Administration has limited options in terms of response. The current mandate, stipulated under a series of 1990 and 1991 U.N. resolutions, is limited to issues related to the defense of Kuwait. It does not provide significant latitude for the alliance to move within Iraq or against Baghdad for internal reasons.

Allocations for U.S. intelligence operations against Hussein's regime have more than quadrupled since the Persian Gulf War ended in 1991. Yet none of the opposition groups from any of Iraq's three major ethnic or religious groups -- northern Kurds, the ruling Sunni Muslim minority or the majority southern Shiites -- is considered to be strong enough to oust Hussein, experts said.

Yet conditions inside Iraq appear to have crossed a threshold, they added.

Inflation was estimated at a staggering 24,000%, and child mortality has quadrupled since the war, Iraqi officials said this summer. The economy generally has deteriorated so badly that the government recently was forced to halve rations originally intended to appease the public because of widespread shortages and soaring prices of basic foodstuffs.

A senior U.S. official said that the shifting circumstances have created "an aura of desperation." But Hussein is a master at using dramatic techniques to mobilize domestic opinion and to manipulate the international community, the official said.

Yet the latest chapter in the Iraqi drama appears to have backfired on Hussein.

Until the crisis, Russia, France, China and Turkey -- three members of the U.N. Security Council and a critical Iraqi neighbor -- favored easing the sanctions after a six-month testing of Iraq's compliance with U.N. monitoring of its weapons of mass destruction, which went into effect Monday. But the United States opposed easing the sanctions, saying that Hussein had not complied with other conditions, such as recognizing Kuwait's borders.

The result of Hussein's troop movements is that those who favored the easing are now more likely to support the call by the United States and Britain for continuing the sanctions, the analysts said. While the coalition recently looked like it might be fracturing over the sanctions issue, it now appears likely once again to unite.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

10.

(92402)<DOCNO> LA100994-0279

SANCTIONS PUT SQUEEZE ON IRAQ ECONOMY

Behind Iraqi leader Saddam Hussein's latest act of military bluster lies the harsh reality of an economy imploding under the weight of international sanctions.

Analysts and economists said Saturday that Baghdad's complaints about the sanctions reflect the very real hardships that four years of restrictions have placed on the country's oil-driven economy.

Iraq's top diplomat, Deputy Prime Minister Tarik Aziz, said in a U.N. speech Friday that the sanctions are leading to malnutrition, inadequate medical care and widespread suffering.

Conditions appear to have reached a critical level in recent months, analysts said, which may explain Hussein's decision to move now.

"I think there is no doubt the sanctions have had an impact on the general population of Iraq," said James Placke, a Middle East analyst with Cambridge Energy Research, an oil industry consulting firm. "The regime has taken steps to protect its security apparatus and to a lesser extent its military from the brunt of the sanctions, but it is clear there has been some impact there as well." Iraq's inflation rate is now more than 1,000% annually, and its currency is virtually worthless. Government-imposed monthly rations for rice, sugar, cooking oil and flour have been halved in the last few weeks, and hunger and starvation are now said to be commonplace, according to international observers who have visited Iraq recently.

There are reports that as many as 500 Iraqi soldiers are deserting from Iraq's 400,000-strong military each month, and there are signs that much of the country's remaining military equipment is only marginally serviceable, Iraqi dissidents have said.

"The military is finding it very difficult to obtain spare parts for its equipment, and it is clear that its readiness is nowhere near the levels it achieved" before the 1991 Persian Gulf War, Placke said.

Indeed, every industry in the country faces critical shortages of parts and equipment, and the average Iraqi household is being forced to make do despite shortages of almost every necessity.

The deprivation has led to a rising tide of violent crime, much of which is reportedly being committed by members of Hussein's ruling clique, desperate to survive the crisis, say dissidents and Western journalists reporting from Baghdad.

The deterioration can be traced, in large measure, to Iraq's loss of its main source of foreign revenue: oil exports.

Before the Persian Gulf War, Iraq was producing as much as 3.3 million barrels of oil per day, the bulk of which went for export. Today, analysts believe that only a small trickle of Iraqi oil is making its way out of the country through cracks in the U.N. embargo, mostly to neighboring Jordan.

Hussein is reportedly so desperate that he has invited foreign oil companies to open potentially lucrative negotiations for future drilling and exploration, hoping to get the industry poised for rebirth after the sanctions are lifted.

As the sanctions bite into Iraqi society, they offer one of the few instances in recent history when international economic restrictions have had a significant impact on an aggressor nation.

Recent efforts to use economic sanctions against Haiti and the rump Yugoslavia, for example, did not bring about U.S. or U.N. policy objectives.

The sanctions' failure in Haiti to force out a military junta created conditions that led the U.S. to intervene to reinstate exiled President Jean-Bertrand Aristide.

But the impact of sanctions on Iraq shows that "it is dangerous to generalize about the effectiveness of sanctions," noted Jeffrey Sachs, an international economist at Harvard University.

"Sometimes they work, sometimes they don't. It all depends on the context, on the commitment of the international community to adhere to them and on what you are trying to accomplish with the sanctions," Sachs said. "But just because they were unsuccessful in Haiti doesn't mean you will see the same result in Iraq." On the other hand, the sanctions against Iraq certainly would not have struck so deep if they had not been preceded by the devastating 1991 military defeat of Hussein's army.

And even with that war as a backdrop, the sanctions have still not yet achieved their primary purpose of forcing Hussein's regime to alter its behavior and comply with all of the U.N. resolutions imposed after the war.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

Appendix B – Relevance Feedback Questionnaire

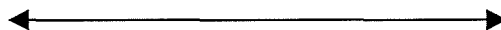
Em uma escala de 1 a 5, como você classifica a dificuldade da tarefa para este tópico (onde 1 significa “muito fácil” e 5 significa “muito difícil”)?

Muito Fácil 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Muito Difícil



Em uma escala de 1 a 5, o quão confiante você está quanto à precisão dos julgamentos que fez para este tópico (onde 1 significa “muito confiante” e 5 significa “muito inseguro”)?

Muito Confiante 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Muito Inseguro



Como você considera a sua familiaridade com o tópico da consulta (onde 1 significa “especialista neste assunto” e 5 significa “não tenho nenhum conhecimento do assunto”)?

Especialista 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Nenhum conhecimento



O regresso de Solzhenitsyn

Encontrar documentos que descrevam o regresso do vencedor do prémio Nobel de Literatura, Solzhenitsyn, à Rússia.

Documentos relevantes que relatem as motivações e a época do regresso de Solzhenitsyn à Rússia. Também poderão mencionar as motivações da sua emigração para os Estados Unidos.

1.

(72138)<DOCNO> LA071994-0037

THE WEEK AHEAD: MOSCOW; SOLZHENITSYN WINDUP

Alexander I. Solzhenitsyn, the Nobel-prize winning writer, completing a nearly two-month tour across his native Russia, is scheduled to arrive in Moscow on Thursday.

Russia's greatest living writer is likely to be coming home cranky. When he returned to Vladivostok on May 27 after two decades in exile, he complained of the intrusive media that greeted him. Later, as he made his way across Russia, he complained that his trip was not being given the serious press coverage it deserved.

However, Solzhenitsyn has agreed to address the new Russian Duma, or lower house of Parliament, next fall.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

2.

(18083)<DOCNO> LA102594-0045

THE WEEK AHEAD: MOSCOW; AN AUTHOR SPEAKS

Alexander I. Solzhenitsyn, the Nobel laureate who was once the scourge of the Soviet Union and is now a critic of the Russian government, will address Parliament members Friday.⁸ The Duma, the lower house of Parliament, invited the author to speak of his return after two decades in exile in the United States. Solzhenitsyn landed in Vladivostok on May 27 and embarked on a long cross-country trip to Moscow to rediscover his homeland.

He has argued that Russia has lost its spiritual bearings and is taking the wrong road away from Communism.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

3.

(58422)<DOCNO> LA052494-0107

THE WEEK AHEAD: MOSCOW; END OF AN EXILE

Alexander I. Solzhenitsyn, Russia's best known author, is due to end 20 years of exile Friday with a flight across the Pacific to the Far Eastern port city of Vladivostok, where he will start a cross-country journey to Moscow and a new home. Solzhenitsyn flew from his homeland in handcuffs with a KGB escort in 1974 after his writings angered the Soviet regime. Stripped of his Soviet citizenship, he and his family lived two years in Geneva and the last 18 years in Cavendish, Vt.

His citizenship was restored in 1990 and Russian President Boris N. Yeltsin invited him to return to post-Soviet Russia. But Solzhenitsyn remained in the United States to finish writing "The Red Wheel," his epic about Russia on the eve of the 1917 Revolution.

Em relação ao tópico, este documento é:

- ☐ relevante
☐ irrelevante
☐ não sei

4.

(97370)<DOCNO> LA102994-0154

SOLZHENITSYN RIPS ALMOST EVERYONE IN PARLIAMENT SPEECH

Alexander I. Solzhenitsyn thundered into Russian politics Friday with his first speech to Parliament -- a blistering dissection of post-Soviet life in which democracy remains just a dream.

Lawmakers greeted the gray-bearded writer warmly and applauded often early in his 50-minute speech. But they cooled and applause grew sparse as he turned his ire on virtually every political current, from market reformers to Communists to the leaders of restive regions and republics.

The 75-year-old writer, who crossed Russia by train last summer after 20 years in exile, cast himself Friday as a messenger from the suffering people.

But some in the audience of the Duma, Parliament's lower house, wondered whether the Nobel laureate wasn't as out of touch with the Russian people as the politicians he condemned.

Solzhenitsyn ran through the painful changes that have accompanied Russia's post-Soviet turn toward capitalism: crime, inflation, industrial decline, poor health, the dominance of the U.S. dollar.

"Russia is emerging from communism on the most twisted, painful and awkward path," he said.

Solzhenitsyn shouted often and gesticulated dramatically. The country, he argued, is still run by Soviet "nomenklatura turncoats disguised as democrats." "We must admit honestly, this is not a democracy. Now we have an oligarchy, since the power belongs to a limited number of people," he said.

Em relação ao tópico, este documento é:

- ☐ relevante
☐ irrelevante
☐ não sei

5.

(12359)<DOCNO> LA072294-0068

SOLZHENITSYN TRIP OFFERS GRIM PICTURE OF RUSSIA; REFORMS: CITIZENS ARE SUFFERING, WRITER SAYS AFTER CROSS-COUNTRY TREK TO GAUGE NATION'S PROBLEMS.

After a 20-year exile in the West and an eight-week rail trek across Russia, Alexander I. Solzhenitsyn returned to Moscow on Thursday night with a litany of complaints from his compatriots and a grim verdict that this country and its reforms are "in big, heavy and multifaceted trouble." "Nobody expected that the way out of communism would be painless, but nobody expected it would be so painful," Russia's greatest living writer declared. "We have tried to break away from communism, but again and again we have taken the clumsiest, hardest and most twisted path.

"Everywhere, you can hear moans that the state is, again, not fulfilling its obligations before its citizens," he added.

His harsh judgment, delivered to about 1,000 people in the rain outside Moscow's Yaroslavl train station, put the 75-year-old sage squarely in the struggle to change Russia's direction. It marked a new challenge for Russian President Boris N. Yeltsin, who in recent months has gained a measure of political and economic stability.

Having spoken with thousands of people on a 5,900-mile journey of rediscovery from Russia's Far East, through Siberia, the Ural Mountains and the Volga River valley, Solzhenitsyn brought added authority to complaints he had uttered from exile that the Yeltsin government's free-market reforms are helping to impoverish this country and corrupt its values. "I wrote down everything they said . . . their advice, requests, pleading, insisting . . . and I will bring all this to the ears of those who have power and influence," Solzhenitsyn said upon arrival.

Then he ticked off their woes: Farmers are being ripped off by middlemen, he said. Huge failing factories are dragging down the welfare of entire cities.

Doctors and teachers are working "almost without reward." The 25 million ethnic Russians who live in neighboring countries are being ignored. "A snake of crime is threatening to strangle all of us," he said.

But his most damning appraisal, which drew the loudest applause, mocked Yeltsin's claim to be Russia's first democratic leader. "People in our country are not masters of their fate," he said. "We cannot say that democracy is in power. There is no democracy here." Russians will be hearing all this again when Solzhenitsyn addresses Parliament after its summer recess. Moving quickly to get Yeltsin out of the firing line, his spokesman suggested Thursday that the author meet soon with the president.

Other politicians of varying stripes lined up to welcome the writer and court him as an ally. Sergei Barburin, a right-wing nationalist, predicted that Solzhenitsyn will side with the "patriotic opposition." Moscow Mayor Yuri M.

Luzhkov, who has presidential ambitions, introduced Solzhenitsyn at the rail station as "the man who best knows Russia from within." Asked from the crowd whether he will run for president, Solzhenitsyn repeated his vow to shun public office. "But I will use all the possibilities to speak and exert influence," he added. "That is my role." Some politicians have suggested that the conservative anti-Communist writer, who has refrained from criticizing Yeltsin by name, will bolster the anti-Yeltsin ranks while moderating their extremism. Apparently sensing his isolation, neo-fascist politician Vladimir V. Zhirinovskiy, whom Solzhenitsyn has labeled a "crude caricature of (a) patriot," said Thursday that the writer should "go back where he came from." Solzhenitsyn has lived 18 of the last 20 years in Vermont. Having spent eight years in Josef Stalin's prison camps and exposed their cruelties in his powerful writings, he was arrested by Soviet authorities in 1974 and stripped of his citizenship.

His east-to-west trek across Russia began May 27 in the Pacific port city of Vladivostok, where he hitched his private luxury rail car to the Trans-Siberian Express. He unhitched it in cities along the

way, staying for days at a time and packing auditoriums and public squares with admirers, many of whom welcomed him as a prophet.

"Seeing Solzhenitsyn is as cleansing as going to a beautiful temple," said Yuri Bakhtin, 35, a businessman who came to hear the writer speak. "He is a living legend. . . . Let's see if he will change us and our ways or if we will change him." Solzhenitsyn's conversion from literary hermit in Vermont to touring celebrity in Russia can only enhance his influence. Having finished his biggest work, a four-volume history of Russia on the eve of the 1917 Bolshevik Revolution, he has returned to a nation that, to his surprise, is more eager to hear his words than to read them. In one poll last month, 44% of those questioned had not read one of his books.

Still, the author said he had "met lots of healthy souls, curious, active, searching minds, often in confusion, not knowing how to direct themselves." In his lone note of optimism, he said Russians "will manage to get out of this pit" but only with "high responsibility at the top and great effort from beneath." Sergei L. Loiko of The Times' Moscow Bureau contributed to this report.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

6.

(8830)<DOCNO> LA052894-0147

SOLZHENITSYN RETURNS FROM 20-YEAR EXILE

Solemn and prophetic, Alexander I. Solzhenitsyn ended his 20-year exile Friday with an appeal to people across Russia to seize the initiative in directing their troubled country's post-Communist rebirth. "I know that I am coming to a Russia torn apart, discouraged, stunned, altered beyond recognition, convulsively searching for itself, for its true identity," the country's greatest living author told an evening homecoming rally in this Pacific port city, his starting point for a Trans-Siberian rail journey to rediscover his country and compatriots.

"I would like, after these meetings, to help you and search together with you for sure ways to get out of our 75-year quagmire," added Solzhenitsyn, who exposed his Soviet tormentors with powerful writings that earned him the Nobel Prize in 1970 -- and expulsion from his homeland Feb. 14, 1974.

About 2,000 people waited three hours on a gray, blustery evening in a seaside plaza for his dramatic return on an Air Alaska commercial flight from the United States. They applauded when his brief speech ended with the words, "I bow to you." "We bow to you!" someone shouted.

Then, answering a question from the audience, the author said that Russia's revival after seven decades of Soviet rule and two years of tumultuous, divisive and stalled democratic reforms would be "difficult, not soon," and would ultimately depend on Russians taking responsibility for their own fate.

"We have been through other difficult times," he said, recalling the Time of Troubles in the 17th Century, when resistance from the hinterlands freed Moscow from Polish rule. "What great initiative we had then! "In reality, our fate is in our hands, starting with every step of life . . .

the little choices taken at the grass-roots level," he said, warming to the populist theme of his journey. "The turn life takes depends on us, not from the decisions bestowed on us from somewhere on high, somewhere in Moscow." "That's true!" called another voice from the crowd.

Judging from his remarks and his ambitious travel plans, the bearded, craggy-faced author intends, at age 75, to play a guiding role in Russian life -- even though he has forsworn seeking elective or appointed office. His voice could provide a rallying cry for regional and local leaders seeking to wrest more autonomy from the central government of President Boris N. Yeltsin.

In an interview with Russia's Itar-Tass news agency, Solzhenitsyn was more categorical about his mission and blunter in his criticism of Moscow, which he said "is leading a privileged life compared with the provinces." He said he was returning home east to west because "to begin with Moscow means to lock oneself in a concrete box." "My literary task is fulfilled," he added. "Now I will have no time to write.

It's time to get down to the hard work of rebuilding and reviving Russia." Many who turned out to see Solzhenitsyn, such as Galina N. Petrovna, said they are ready to follow his prescriptions as those of a sage.

Petrovna, a graying, bright-eyed woman of 69, came to the square bedecked with Soviet medals earned for underground resistance in a Nazi prison camp. But she was a reluctant hero, having lost her father in the Stalin regime's political executions.

"I have been waiting for this moment all my life," she said at the rally for Solzhenitsyn. "Here is a great man who can show us how to rebuild our country.

If we had worked like he does, by the call of our souls and hearts, we would live now in a different, much happier country." Anticipating the writer's long journey home in a front-page headline, the Moscow newspaper *Nezavisimaya Gazeta* declared, "Solzhenitsyn, like the sun, rises in the east." The trip began Wednesday in an Oldsmobile station wagon, lumbering down the unpaved driveway from the hilltop country home in Cavendish, Vt., where Solzhenitsyn spent the last 18 years writing in near-seclusion.

With a newly issued Russian passport, the author flew to Anchorage, Alaska, with his wife, Natalia, and Stephan, 20, the youngest of their three sons. They left behind the middle son, Ignat, 21, and Natalia's mother, Yekaterina Svetlova, who will join them later. The eldest son, Yermolai, 23, flew here, 5,700 miles from Moscow, ahead of his parents to help arrange the overland journey to the capital, which he called "the greatest road trip you could do." Air Alaska's first touchdown in Russia came at Magadan, once a main receiving point for those destined for the Soviet gulag prison camp system. Solzhenitsyn, who spent eight years in the gulag and exposed its evils in his best-known works, once described Magadan as the site containing the most human bones on Earth.

Stepping off the jet, Solzhenitsyn stooped, touched the tarmac with both hands and crossed himself. "Today, in the heat of political change, those millions of victims are too lightly forgotten, both by those who were not touched by that annihilation and even more so by those who were responsible for it," he said.

"Under ancient Christian tradition, land where innocent victims are buried becomes holy. We shall consider it so, in the hope that the light of Russia's coming recovery will reach (this) region." Landing in Vladivostok, he accepted an offering of bread and salt -- a Russian symbol of hospitality -- and raised the loaf to his lips. He also kissed Igor P. Lebedinets, the acting territorial governor, who postponed an official visit to South Korea to organize a welcome. The city's Russian Orthodox bishop was on hand to give his blessing.

"All the best people are leaving Russia, and Solzhenitsyn is the one coming back," said Antonina N. Detyareva, 39, who was selling dried squid and imported apples just off the puddled airport parking lot. "Maybe that will set a good example." The writer returned a controversial figure, however, scorned by many who miss the old order and condemn him for helping destroy it.

"He'll get no warm welcome from me," said Vladimir A. Gornoslal, 45, a taxi driver at the airport. "I don't like traitors." Arriving in the seaside square, the writer found it still dominated by a Stalinesque statue, four stories high, of a proletarian soldier-hero. "To those who struggled for Soviet power in the years 1917-1922," the caption reads.

But this raucous city of 700,000, closed to outsiders until 1992, is a leading arena of post-Soviet transformations, good and bad. Hundreds of joint ventures with Japanese, South Korean, Chinese and U.S. firms are under way. But crime and corruption are flourishing, and the gap between newly rich and newly impoverished is growing. City services are crumbling.

Solzhenitsyn turned down offers by the local authorities to stay in more luxurious quarters and chose the Vladivostok, a \$90-per-night hotel that, along with most of the city these days, has no hot water. Sergei Loiko, a reporter in the Times Moscow Bureau, contributed to this report.

Em relação ao tópico, este documento é:

- ☐ relevante
☐ irrelevante
☐ não sei

7.

(63755)<DOCNO> LA061594-0027

BOOK REVIEW / HISTORY; AS YELTSIN SEES IT -- INCLUDING REVISIONS; THE STRUGGLE FOR RUSSIA; BY BORIS YELTSIN, TRANSLATED BY CATHERINE A. FITZPATRICK (TIMES BOOKS: \$25, 316 PAGES) Now that Solzhenitsyn has renounced literature in favor of politics, is there room in the Russian literary pantheon for politician-turned-author Boris Yeltsin? * Uh, no.

Yeltsin's second and latest book, "The Struggle for Russia," is a political memoir fashioned out of excerpts from a journal that he apparently managed to keep during a tumultuous three-year period that included the failed coup of August, 1991; the final disintegration of the Soviet Union; the shoot-out at the Russian parliament, and miscellaneous other crises.

"Whether people like it or not," the first elected president of Russia muses out loud, "they must figure out who this Yeltsin is -- what kind of person is he, what does he think, where is he taking us?" While Yeltsin's book is, at moments, a worthy account of contemporary Russian politics from an insider's point of view, Yeltsin neither aspires to nor achieves much intimacy in his reflections or much grandeur in his prose. Rather, he presents himself as a hard-working politician who is really a just-plain-Ivan kind of guy.

Now and then, Yeltsin allows us to glimpse the more private facets of his life: "I suffer from insomnia," he explains, and he confesses to stage fright: "When I have to go on TV, I sweat bullets." Yeltsin is far more interested in explaining (or justifying) himself in the eyes of history. He often refers to himself in the third person as "Yeltsin" or, more often, "the president." And he invariably styles himself as a democratic activist -- his attack on Mikhail Gorbachev in the last days of the Soviet Union, for example, is explained as a matter of principle, not politics.

"Unlike most democrats, I surmised that the threat of dictatorship came not only from 'Gorby's' entourage but also from Gorbachev himself," Yeltsin declares. "And that was truly terrible." Only rarely does Yeltsin give us a truly startling insight into his own political destiny, as when he ponders the yellowing papers in a secret police file on the arrest of his father during the worst years of the Stalinist terror: "I was only three years old at the time of my father's arrest," he recalls, "but I remember to this day all the horror and fear." More often, Yeltsin contents himself with dishing the inside story on Kremlin politics in the final days of the Soviet Union and the stormiest moments of its tumultuous aftermath.

Yeltsin also rises to his own defense in explaining his trepidation in the days leading up to the assault on the White House. Condemning the insurrectionists as "neo-Nazis, bandits and outright fascists," he insists that he acted courageously and decisively to crush the threat to the infant Russian democracy.

The publisher is candid enough to advise us that "in a few places, Yeltsin himself has exercised the author's prerogative to amend his observations and place them in perspective." In other words, Yeltsin has rewritten his journal entries so that they support his claim to a place in world history.

"Something amazing happened," he writes about the collapse of communism.

"Overnight, a new Russia -- 'Boris Yeltsin's Russia' -- took the place of the Soviet Union in international politics." We do not yet know whether Yeltsin, like Gorbachev, will end up on the dust heap of history. But, in the meantime, Yeltsin is hoping that readers in Russia and America, too, are still interested in knowing he thinks of himself and his curious fate as "the first freely chosen president of Russia."

Em relação ao tópico, este documento é:

- ☐ relevante
☐ irrelevante
☐ não sei

8.

(57835)<DOCNO> LA052194-0202

SOLZHENITSYN HEADS HOME TO WRITE RUSSIA'S WRONGS; CULTURE: DISSIDENT AUTHOR FACES UNEASY HOMECOMING IN A LAND CAUGHT IN A PAINFUL POST-SOVIET REBIRTH.

The newest estate on a bluff over the Moscow River is an imposing red brick dacha with two stories, two wings and a high green fence secluding its 10 pine-shaded acres. From the outside, it looks worthy of the great man who plans to spend his final years here.

Alexander I. Solzhenitsyn, the dissident writer who exposed and outlasted his Soviet tormentors, is returning home next Friday after 20 years in exile to rediscover his beloved Russia and offer his wisdom to help guide its painful post-Communist rebirth.

But before he can weigh in on such issues as the gulf between Russia's rich and poor, the grip of organized crime, the spread of Western consumer culture and the specter of fascism, he faces an irritating little problem right under his new doorstep.

The basement leaks.

In his first lesson about the new Russia, Solzhenitsyn has fallen victim to the capitalist version of an old Soviet problem -- khaltura , or shoddy workmanship. As often happens here, the author did not have a contract and the architect refused to do it over for free.

The affront may be trifling compared to Solzhenitsyn's eight-year imprisonment in Josef Stalin's gulag and expulsion by Leonid I. Brezhnev's KGB, but it is just one sign that the 75-year-old sage is in for an uneasy homecoming. Apart from the trials of daily life, he returns as a controversial figure.

Many Russians look forward to the reunion; students born after he left are even writing essays anticipating it. Older people say they've longed for a moral leader of his stature since the death of fellow dissident and Nobel-winning physicist Andrei D. Sakharov in 1989. Politicians across the spectrum are poised to pounce on his every word as proof of their version of the truth.

But some Russians curse Solzhenitsyn and his writings as a destroyer of an old order they miss. Others say he may be returning too late to shape the new one.

In a country changing so quickly and seemingly so adrift, many Russians ask, What difference can any great artist or prophet make? "He doesn't have a real understanding of the country he's coming back to," says Mikhail Y. Shvydkoi, Russia's deputy culture minister. "Perhaps this is because people who live in this country don't have a real understanding of it either." Some are simply afraid for him. Mail from friends in Russia is running 9 to 1 against his return, according to his wife, Natalia.

"They are saying it's a terrible situation now," she reports. "Don't go. You are very valuable for Russia, but you won't be safe." But Solzhenitsyn, who settled with his wife, mother-in-law and three sons in Cavendish, Vt., in 1976, has been planning his return for three years since finishing "The Red Wheel," an epic history of Russia on the eve of the 1917 Revolution.

His citizenship, stripped away by the Soviet regime, was restored in 1990.

Russian President Boris N. Yeltsin phoned two years later with a personal appeal to come home.

Shvydkoi says the government plans "a very honorable ceremony." Mindful of his preference for rural surroundings, Moscow's city government last year gave Solzhenitsyn land use rights here on the western outskirts of the capital, the "dachaland" of past and present Kremlin elites. It also sold him the buildings on the 10-acre property, which originally belonged to a Stalinist henchman named Lazar Kaganovich.

The Solzhenitsyns had the old house torn down and a new one built at their own expense. One of the two wings will be the family's living quarters, the other a workplace and repository of the author's extensive library and archives, which will be open to researchers.

As the family began packing in Vermont last winter, plans to finish the house this spring went awry.

Snowfall on the roof turned to ice, widened the seams between the copper tiles and melted into the house when the heating system came on. More water leaked into the basement from a poorly insulated outdoor porch.

"It was a disaster," says Mrs. Solzhenitsyn, who came to Moscow to clean up.

She replaced the architect and the roof, then halted all other work, leaving the basement to be repaired this summer.

"The house is only walls and roof now," she told *The Times* in a telephone interview from Vermont. "There's nothing inside, not even a floor. . . . We decided it was impossible to continue the work until we live in Russia. People take advantage. You have to be around." Until the house is ready, the family will live on Moscow's Plyushchikha Street in a five-room apartment that Mrs. Solzhenitsyn bought two years ago.

She says her husband will spend much of the summer traveling around Russia "meeting lots of people and listening to them about life today," then start writing short stories and playing "some role as a public figure" outside politics.

"He won't recognize the country when he comes back," says Irina Nefedyeva, a 35-year-old teacher, as she shopped on Plyushchikha the other day. "Everything is changing, but it's hard to say in which direction." Along the street, a Russian Orthodox church closed by the Communists has reopened, but a textile plant is struggling, and two bakeries have closed. To eke out a profit, a tiny food shop also sells Japanese-made TV sets and VCRs.

The newest store specializes in American car accessories.

A lot of what the author disdains as "squalid mass culture" from the West -- novelized versions of the Charles Bronson movie "Death Wish," an Italian TV series "Octopus," a Mexican soap opera "Simply Maria" -- is available at the neighborhood bookstore, which didn't bother reordering Solzhenitsyn's "Gulag Archipelago" and "One Day in the Life of Ivan Denisovich" when they ran out a few months ago.

"To put it mildly, Solzhenitsyn's books are not as popular now" as when they were banned under the Soviet regime, says Alexandra V. Nezhikova, who works in the store. "When there are none of his books for sale, nobody asks for them." Even some of Russia's intelligentsia admit having trouble staying awake with "The Red Wheel." The political and cultural elite are far more interested in how Solzhenitsyn judges 1994.

Solzhenitsyn laid out some of his views this month in an interview for *Forbes* magazine that was reprinted by the Russian newspaper *Izvestia* and found some common ground with Communists and radical nationalists.

For example, he suggested that Yeltsin was wrong to privatize farmland and is not adequately defending the rights of ethnic Russians in the former Soviet republics. He also warned that some Americans want to "weaken Russia." Democratic reformers, who have long suspected Solzhenitsyn of autocratic instincts, reacted with fear that he will exert a reactionary influence on Yeltsin. For his part, Communist Party leader Gennady A. Zyuganov sounds encouraged.

"He will come back and he will not be happy at all with the picture unveiled before his eyes," Zyuganov says. "We will gratefully accept his words of truth." Those inspired by Solzhenitsyn to

fight the Communist regime say the most they expect of him now is to stay above the political fray and somehow make it more civil.

Lev Timofeev, 57, a dissident jailed in the mid-1980s, credits the author with giving his generation "the language of freedom." "The most acute problem now is how to find a language of social accord," Timofeev says. "When Yeltsin proclaims this, a great number of people dismiss it as a political trick. But if such an idea is proclaimed by Solzhenitsyn, that's absolutely another thing. I hope that Solzhenitsyn will create this new language." Russian literary figures of earlier eras -- Dostoevsky and Tolstoy under the czars, Pasternak under the Bolsheviks -- assumed such roles as public figures and moral critics. But they were present to witness the changes of their times for themselves.

"Solzhenitsyn's authority is weakened by his distance from events," says Benedict M. Sarnov, a leading Russian literary critic. "He is counting on being a guru, but I think he will be disappointed. I would advise him not to come back. . . . What Russia needs now is not prophets but pragmatists." The debate echoes in Russian high schools, where Solzhenitsyn's works are taught alongside those of Pushkin, Tolstoy, Gogol, Bulgakov and Dostoevsky.

At School No. 1289 in Moscow, 11th graders say it's cool to throw around Solzhenitsyn's name, even if they don't actually read him or can't remember what gulag means. One classroom discussion turned into a spirited argument between two 17-year-olds, Timofei Kiselev and Yelena Sayenko.

Timofei: "Why is Solzhenitsyn coming back? What can he do here? We have a very unstable economic and political situation. The country is not ready to accept him." Yelena: "We should accept him. He's a genius." Timofei: "Sure, but what can he do about our economy?" Yelena: "He can help us understand each other, to live in peace." Timofei: "Not now. Maybe in the future." Yelena: "He'll be dead." Timofei: "His books will always be around. But the country doesn't need him now. We need first of all not a good writer but good government, economic stability." Burdened by daily hardships, some Russians view Solzhenitsyn's return next week with a what's-in-it-for-me attitude, either cynical indifference or a naive expectation that it will improve their material lot.

Hope is high among the modest villagers who live outside the fence around Solzhenitsyn's unfinished dacha. Relieved that it was the great writer and not some Mafioso or Kremlin big shot moving in, they are wondering whether he might get the neighborhood some gas heating or at least a sewage system.

"We are looking forward to meeting him and talking to him," says Viktor S.

Strunkin, a 50-year-old radio mechanic, stopping on the muddy road by a garbage dump 100 yards from Solzhenitsyn's gate.

"Maybe the authorities will remove this awful dump," he adds. "Maybe they will pave this road. I can't imagine that Solzhenitsyn will love having a garbage dump right under his nose."

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

9.

(65962)<DOCNO> LA062394-0406

LAWMAKERS REJECT YELTSIN'S ANTI-CRIME PLAN AS TOO HARSH Russia's Parliament sent President Boris N. Yeltsin a resounding nyet Wednesday, voting overwhelmingly to reject a presidential anti-crime plan that foes say would wipe out many hard-won new constitutional rights.

Despite a crime wave that has terrorized and enraged many Russians, lawmakers protested that Yeltsin's harsh cure smelled too much of the odious practices of Soviet days.

His June 12 decree would allow police to detain suspects for 30 days without charges; search offices, cars and passengers without warrants, and inspect bank records of those suspected of organized crime without obtaining a court order.

It would also make evidence gathered by wiretapping admissible in court.

"This is exactly how it all began in the '30s," said radical nationalist lawmaker Sergei N. Baburin, referring to Soviet dictator Josef Stalin's reign of terror. "We all know what this finally led to." The Stalin-era terror was heralded by a 1934 law-and-order decree that made it easier to convict criminals. The campaign is believed to have claimed the lives of at least 20 million people.

On Wednesday, outnumbered democrats in the 6-month-old Duma, the lower house, found themselves -- to their astonishment -- agreeing with Communists and nationalists in the 279-10 vote. Parliament lacks the power to override a presidential decree, but Yeltsin's staunchest allies voted to ask him to suspend the anti-crime measures.

The rebuff came even as Yeltsin scored big international gains. In Washington, Russian officials announced that a Clinton-Yeltsin summit will be held in September; in Brussels, Russia at last signed the North Atlantic Treaty Organization's "Partnership for Peace" accord -- with a side treaty that gave the fallen superpower bragging rights to much of its lost prestige.

Meanwhile, in what appeared to be a trial balloon, Vladimir F. Shumeiko, leader of Parliament's upper house, told various news organizations that Yeltsin and parliamentary leaders had agreed "in principle" to postpone the 1996 elections for two years.

Ivan Rybkin, leader of the lower house, initially said he liked the idea but on Wednesday told Russian television that it should be treated with caution.

The constitution that Yeltsin begged voters to adopt six months ago calls for presidential and parliamentary elections in 1996.

Shumeiko has supported Yeltsin's most unpopular decisions. When Yeltsin dissolved the old Supreme Soviet in September and promised to run for reelection this month -- then changed his mind after his foes were defeated -- Shumeiko took to the hustings to defend the president.

There was no immediate reaction to Shumeiko's statements from the Duma, which was preoccupied with forcing Yeltsin to revise his crime plan.

Former dissident Sergei Kovalyov, chairman of the Human Rights Committee, called the president's decree an intolerable violation of the civil rights guaranteed by the new constitution. Kovalyov, who was sent to labor camp for defending Andrei D. Sakharov, the late Nobel laureate and human rights activist, said Russia has a bad tradition of letting doctors, not patients, write laws on health care -- and police, not citizens, draft criminal codes.

He cited an instance, in the Volga republic of Bashkortostan, in which police were recently permitted to detain suspects for 30 days after a state official was slain.

"Five hundred suspects were detained in a town of 35,000," Kovalyov said.

"Three of the suspects tried to commit suicide, and two quite independently of each other pleaded guilty." But Nikolai I. Travkin, leader of the conservative Democratic Party of Russia, called Kovalyov one of "God's fools." "They would let half the country be shot in the street but not a single criminal's human rights may be violated," he said. "Today we will not conquer crime by normal, civilized methods. It is overwhelming us." Yeltsin's decree took a drubbing in Moscow's most influential newspapers. One of its only backers was Alexander I. Solzhenitsyn, the famed, Nobel Prize-winning author who returned to Russia last month from exile in Vermont and is now lamenting the moral decline of his homeland.

Neo-fascist lawmaker Vladimir V. Zhirinovsky, predictably, denounced the Yeltsin decree as insufficiently tough. He said he would introduce his own Draconian plan Friday.

Yeltsin's promise to crack down on "criminal filth" followed a bloody spring of bombings, killings, kidnaping, hostage-taking and hijacking, as well as street crime -- which has grown all too common here.

One Yeltsin aide warned lawmakers not to discredit themselves by appearing to be soft on crime; both sides promised to seek compromise. The Kremlin and Duma are each at work revising the unwieldy Soviet-era criminal code.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

10.

(60004)<DOCNO> LA052994-0317

SOLZHENITSYN OPENS A NEW CHAPTER OF HIS LIFE: CELEBRITY On the first day of a cross-country tour to rediscover his Russian homeland, Alexander I. Solzhenitsyn visited a hospital Saturday. As the famous author stepped into an elevator, a small crowd of doctors and a television film crew jammed in with him.

Instead of going up to the director's office on the sixth floor, the overloaded elevator sank part way between the lobby and the basement and stuck there for 15 minutes before a mechanic came to the rescue.

Aside from the hazards of everyday life here, the mishap reflected the quandary of Solzhenitsyn's identity as he returns from two decades of exile to help post-Soviet Russia find itself. Is he an intellectual or a celebrity? Can't he have peace to search for new truths, or must he put up with media stardom and intrusive fans? And how long can the 75-year-old Nobel laureate depend on special protection from a government he criticizes so fiercely? Solzhenitsyn spelled out his controversial views on Russia's ills at a wide-ranging news conference, insisting that there is no real democracy or economic reform here, no cleansing of the Communist legacy and too much imitation of the West. He was equally passionate about his privacy.

"I would prefer not to have 200 photographers crowding around me all the time and shooting me day after day," he said. "I need to talk to the common people .

. . . to learn in detail about life in Russia." To be more specific, he said he resents the fact that somebody in the crowd stepped on his wife's foot after their triumphant landing here Friday on a flight from Alaska.

"When it's enough for each photographer to take two shots, why take 250?" he asked.

After talking with doctors and surgery patients at the Territorial Clinical Hospital, Solzhenitsyn rode to the city's Pacific port, where he was greeted by a navy band playing a lively classical medley, well-wishers pressing for autographs and more cameras.

Strolling down a waterfront aisle of merchants selling food from truck-sized freight containers, the writer acted the part of a modern Rip van Winkle just awakened from a 20-year nap.

"How much is that?" he asked a man selling sausages.

"10,000 rubles," the merchant replied. About \$5 a kilogram.

"I don't get it," Solzhenitsyn said, touching his long beard and appearing amazed. "When I left the country many years ago, prices were very different." "Come again soon, sir," the merchant said with a laugh, "and I'll sell it to you for an even higher price!" Plainclothes police officers locked

arms in a protective ring around Solzhenitsyn and his wife as they walked, while uniformed officers stopped people at random to inspect purses and shopping bags.

"While he's here, we will keep him under guard whether he likes it or not," said Igor P. Lebedinets, the acting territorial governor. "We're mobilizing hundreds of men." Later, Solzhenitsyn was driven in a black government Volga sedan to have lunch aboard one of the two luxury rail cars sent from Moscow by the Railway Ministry for his trans-Siberian journey to the Russian capital. The cars, equipped with a kitchen and Oriental carpets, will carry his entourage of family, friends, chefs, attendants, security police and a British Broadcasting Corp. documentary film crew.

Solzhenitsyn said the 5,700-mile rail journey, which will take several weeks, will allow him to make many stops and to "see things through the window." He said he last saw Siberia from a prison van, on the way to an eight-year term in Stalin's gulag after World War II -- the seminal experience for his powerful writings against the Soviet system.

The lavish official hospitality by the current authorities prompted one Russian journalist to ask Solzhenitsyn about comparisons to Maxim Gorky, the writer brought back from Italian exile under Stalin.

"Gorky came back to serve the regime," Solzhenitsyn replied. "I will never serve the regime, whatever government is in power." But he refrained from criticizing Russian President Boris N. Yeltsin, who sent a congratulatory telegram saying, "Your talent and your experience as a historian and thinker will help us all in reorganizing Russia." The writer did, however, repeat many criticisms he had made abroad about the Yeltsin government's attempts at reform.

For example, he described the 1992 decisions to free up prices and sell off state property without breaking up monopolies as "brainless" steps that had made a country already wrecked by the Communists even poorer.

He called Russia a "pseudo-democracy" that still lacks local self-rule, has failed to punish the crimes of Soviet repression and blindly imitates Western ways unsuited to its culture.

He sneered at the seepage of such foreign terms as "voucher" and "futures" into Russian usage.

"It's not simply the sickness of our language," he declared, "it's the sickness of our soul." Insisting that outsiders must "give Russia a chance to pull together as Russia," he implicitly called for joining the Russian-speaking parts of neighboring Kazakhstan, Belarus and Ukraine with Russia proper.

At the same time, he dismissed Vladimir V. Zhirinovsky, the neo-fascist politician who played up nationalist themes in last year's parliamentary election, as a "caricature of a Russian patriot." While Solzhenitsyn denied any political ambitions, his more measured Russia-first views are likely to strike a powerful chord as he travels and speaks across the country.

"The atmosphere in the hospital changed in one minute after he came," said Yelizaveta B. Pyatina, 58, a patient there. "Everybody began smiling and talking about books, politics, history. We all forgot about our ailments." Some Russian politicians and intellectuals have criticized Solzhenitsyn for coming home too late to shape Russia's development.

The newspaper Moscow Komsomolets reflected that view with a headline: "The old man missed the plane." The "old man" was not amused. When that newspaper's correspondent rose to ask a question at his news conference, Solzhenitsyn chided him for the headline and ordered him to sit down.

"You'll have to wait for the next plane," Solzhenitsyn said.

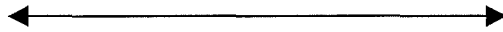
Sergei L. Loiko, a reporter in The Times' Moscow Bureau, contributed to this story.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

Em uma escala de 1 a 5, como você classifica a dificuldade da tarefa para este tópico (onde 1 significa “muito fácil” e 5 significa “muito difícil”)?

Muito Fácil 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Muito Difícil



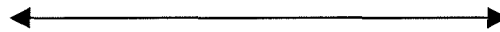
Em uma escala de 1 a 5, o quão confiante você está quanto à precisão dos julgamentos que fez para este tópico (onde 1 significa “muito confiante” e 5 significa “muito inseguro”) ?

Muito Confiante 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Muito Inseguro



Como você considera a sua familiaridade com o tópico da consulta (onde 1 significa “especialista neste assunto” e 5 significa “não tenho nenhum conhecimento do assunto”)?

Especialista 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Nenhum conhecimento



Engenharia Genética

Como Engenharia Genética afeta a cadeia alimentar?

Artigos devem abordar diretamente a introdução da Engenharia Genética e os seus efeitos na cadeia alimentar. Deverão discutir vantagens e desvantagens. Relatórios sobre engenharia biológica no tabaco e engenharia genética humana não são relevantes.

1.

(15661)<DOCNO> LA091594-0261

INOVAÇÃO / MICHAEL SCHRAGE: E EM SUA ESQUERDA, VOCÊ PODE APENAS FAZER PARA FORA O GENE INSIGHTFUL DO COLUMNIST

Os cartographers os mais importantes no mundo que estes dias não trabalham na margem McNally ou o CIA. Não estão muito interessados nas montanhas e em limites nacionais dos rios e deslocar. No fato, não sabem necessariamente muito sobre a geografia tradicional em tudo. O No., os fabricantes influencial do mapa dos 1990s dobra seu comércio nos laboratórios molecular da biologia, fazendo um mapa da geografia dos genes.

Estão criando rapidamente um atlas veritable dos genes do ser humano, da planta e do animal, figurando para fora onde os genes individuais são encontrados e o que . Este está provando ser um dos riscos os mais produtivos no history do cartography: traçar do gene está transformando a biologia, a agricultura e a medicina tão certamente quanto traçando mundo novo a Europa transformada. Este ano sozinho, os investigadores encontraram mais de 45 genes doença-disease-causing novos nos seres humanos -- apenas ontem, os cientistas anunciaram a descoberta do gene que é acreditado para causar a metade de todas as caixas herdadas do cancer de peito.

O ritmo de melhor-do que-gene-um-semana ajudou ao spawn o crescimento de negócio do biotech. Mas que bom é um atlas sem um guidebook? Alas, não há nenhum Zagats ou a guia Michelins para o genome apenas ainda, nenhuma excursões formais dos pontos quentes uncharted sobre chromosomes. Para ajudar retificar este problema, alguns peritos abaixo indicam seus locais favoritos no mapa do gene e explicam onde estão os lugares os mais interessantes a visitar: * " os genes de HOX, " ofertas Edward P. Lewis, professor de Morgan da caça de Thomas da biologia molecular em CalTech. " são a única maneira que um organismo pode fazer uma cabeça, thorax e abdômen. Fazem-na dizendo aos outros genes o que fazer. Essa parte do gene não mudou por sobre 500 milhão anos.

" * " LINHA 1 e LRE1, LRE2 e LRE3, " diz o singer de Maxine, um biólogo molecular que sejam presidente da instituição de Carnegie de Washington e que tenha um laboratório no instituto nacional do cancer. " são os lugares onde as cópias ativas das únicas cópias sabidas de transposable humano -- genes saltando so-called -- são encontradas. São mutagenic quando saltam: isto é, causam mutations e doença, como o dystrophy muscular, quando se movem. Parecem ser ativos em pilhas embryonic adiantadas. São troublemakers, e eu estou particularmente interessado dentro porque estes troublemakers persistiram para assim que por muito tempo no genome por causa de todo o problema que causam.

" * " Rin, " diz William R. Hyatt, vice-presidente da pesquisa e do desenvolvimento para Calgene, companhia que desenvolveu o tomate genetically projetado de FlavrSavr. " é um ponto no locus do tomate e os mutations traçam em lá. É um mutant natural e inibe ripening. Não é assim muito um gene porque é um local. O ponto é, ninguém sabe exatamente o que este gene ou locus são -- exerce este multitude dos efeitos no tomate e através de que mecanismo. Nós temos começado somente recentemente a clone o.

" * " o gene sry do chromosome de Y, " ofertas Robert L. Sinsheimer, um professor emeritus em Santa UC Barbara e um pioneiro nos dias adiantados da pesquisa bacteriana do vírus. " este foi pensado para ser o gene central para o maleness determinar e o femaleness. Há genes também correspondentes no chromosome de X (fêmea). Nós pensamos que nós compreendemos que genes lhe fizeram um menino ou uma menina. Pareceu que desobstruído bonito e agora ele é confused.

" * " os genes responsáveis para o diabetes do tipo 1, " diz o capitalista de risco Kevin Kinsella, que é também executivo principal do therapeutics Jolla-jolla-based la de Sequana. " os locais são os chromosomes 6, 11 e 18 e HLA próximo. Pode haver tanto como outros cinco genes implicaram. Eu sou muito mais interessado agora em como os complexos do gene interagem para produzir uma doença melhor que apenas procurar únicos genes. O problema está encontrando todas aquelas posições múltiplas. Assim eu penso que nós nos estamos movendo em uma era nova da descoberta do gene.

" * " os genes de Synthase do ACUMULADOR, " sugerem David Evans, vice-presidente do desenvolvimento de negócio para a tecnologia da planta do DNA, que foi envolvida em criar o marrom & o tabaco elevado do nicotina de Y-1 de Williamson. " oversee o biosynthesis do etileno. Controlam ripening e rotting. É uma família do gene envolvida na parte dos mecanismos da defesa da planta de encontro à doença. Em certas frutas, como bananas e tomates, há uma associação direta entre o etileno e rotting da fruta. Nós estamos olhando para suspender aquele.

" * " os genes para o receptor de NMDA, " ofertas Daniel Koshland, um biólogo molecular em Berkeley UC e o editor do compartimento da ciência. " é o receptor no cérebro envolvido em como nós gravamos a memória. Eu estou tentando compreender o deslocamento biochemical entre a memória a curto prazo e a memória a longo prazo. Nós estamos procurando o gene para ele. Onde está no chromosome não é que importante neste momento -- mas pode ser. Eu não tenho nenhuma garantias que sua posição é importante mas aquela pode girar para fora para nos dar introspecções importantes. Às vezes a geografia do gene pode ser muito importante.

" * " os conjuntos do gene do histone, " diz Laurence Kedes, diretor do instituto para a medicina genetic em USC. " estes são os genes que fazem as proteínas que revestem os chromosomes. São esses responsáveis para fazer exame das linhas do DNA e comprimi-las em pilhas. Se não, seriam jardas por muito tempo

" * Os investigadores MÉDICOS da DESCOBERTA identificam o gene responsável para alguns cancers de peito. A1

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

2.

(98683)<DOCNO> LA110394-0361

SETE ALIMENTOS MAIS GENETICALLY ALTERADOS PASSAM TESTES DA SEGURANÇA DO FDA

Sete alimentos mais genetically projetados passaram uma inspeção voluntária de segurança, o alimento e a droga quarta-feira dita administração, mas os cientistas questionados se o FDA scrutinizing tais alimentos próxima bastante. " eu sou um pouco incomodado, " Dr. Marion Nestle, um conselheiro do FDA da universidade de york novo, disse a agência. " é como se os cientistas do FDA aceitaram estas avaliações muito complicadas (de segurança) no valor de cara."

Entre os alimentos novos é uma batata produzida por Monsanto que resiste naturalmente o besouro da batata de Colorado. A batata produz um pesticida natural que mate os besouros que fazem exame de uma mordida. O FDA última mola aprovou alimento do biotech da nação o primeiro, o tomate de Flavir-Savr de Calgene Inc., que tem um gene que permita que ripen mais por muito tempo na videira e comece ainda aos supermarkets sem soft de giro.

A agência tem terminado agora inspeções muito mais cursory de segurança da batata e outras de seis plantas geneticamente alteradas, Dr. principal Jim quarta-feira dita Maryanski da segurança do alimento do FDA. Foram aprovados também:

* Três mais tomates, um que trabalha como o Flavr-Savr e os dois com genes alterados que retardam ripening natural dos tomates, assim que eles podem ser escolhidos mais cedo e ripeness optimal do alcance pelo tempo onde batem prateleiras da loja.

* Um squash alterou-se geneticamente para resistir dois vírus mortais.

* O algodão que tolera um herbicida comum, bromoxynil, assim que os fazendeiros poderiam usar o herbicida. O algodão é pretendido produzir o óleo de semente de algodão para a alimentação animal.

* Plantas do soybean que toleram o glyphosate do herbicida. Embora o FDA diga a planta será usada para a alimentação animal, soybeans e o óleo de soybean é consumido por povos. O FDA está preparando-se para exigir que esteja notificado antes de todo o alimento geneticamente projetado esteja introduzido no mercado. Isso dará à agência uma possibilidade decidir-se que tipo da revisão de segurança cada um necessita. " nós simplesmente não temos os recursos para dar cada planta que a revisão intensiva, e nós não vemos a necessidade da saúde pública, " Maryanski dito.

Mas os cientistas advertiram que estes alimentos devem ser prestados atenção próxima para se assegurar de que os povos com allergies não estivessem postos no perigo se, por exemplo, uma proteína da porca fosse posta em soybeans. Alguns ecólogos foram preocupados também, Reuters relatado. Jeremy Rifkin, cabeça da campanha pura do alimento, um grupo oposto à engenharia genética do alimento, dita: " quando você libera uma planta geneticamente projetada no ambiente, os organismos estão vivos -- podem reproduzir, eles mutam da lata e você não pode recordá-los ao laboratório."

Em relação ao tópico, este documento é:

- ☐ relevante
☐ irrelevante
☐ não sei

3.

(15689)<DOCNO> LA091594-0289

MORDIDAS DA NOTÍCIA: O ÚLTIMO UM NO POOL É GALINHA

Muitos de nós acreditam que nosso craving para certos alimentos é assim que profundamente não poderia apenas ser o resultado de como nós fomos trazidos acima, ele tem para ser hereditário. Scientists discordou, e tendido a acreditar o gosto seja devido à cultura ou ao upbringing, mas agora há alguma evidência de um elemento genético.

Os cientistas na universidade de Cincinnati estudaram as preferências do alimento dos jogos dos gêmeos e concluíram que o gosto para determinados alimentos parece ser afetado fortemente pelo genética. Os alimentos que parece ao greve uma nota no pool de gene incluem o suco alaranjado, a galinha, o queijo de casa de campo e os bróculos. BRUCE HENSTELL

Em relação ao tópico, este documento é:

- ☐ relevante
☐ irrelevante
☐ não sei

4.

(15507)<DOCNO> LA091594-0107

A DESCOBERTA DO GENE DO CANCER DE PEITO CHAMOU O AVANÇO PRINCIPAL; MEDICINA: OS INVESTIGADORES ENCONTRAM O DEFEITO QUE CAUSA A METADE DE CASOS HERDADOS. A DETEÇÃO ADIANTADA PODE CONSERVAR MUITAS VIDAS.

Os investigadores de Utá identificaram o gene defeituoso acreditado para causar sobre a metade de todas as caixas herdadas do cancer de peito, uma descoberta que devesse a fazer possível dentro de um ano identificar as 600.000 mulheres americanas estimadas que carregam o gene.

Estas mulheres têm um risco 85% de desenvolver o cancer de peito pela idade 65 e um risco raramente elevado de desenvolver o cancer ovarian. Identificar as mulheres no risco poderia conservar muitas vidas permitindo a monitoração intensiva e a deteção adiantada dos tumores, que aumenta substancialmente a taxa da cura. A descoberta é o resultado de um esforço maciço provocou quatro anos há quando o rei do geneticist Mary-Claire de Berkeley UC mostrou que o gene, chamado geralmente BRCA1, está ficado situado em uma região pequena do chromosome 17, um dos 23 pares dos chromosomes que contêm o blueprint genetic de cada pessoa.

A raça para encontrar o gene foi ganhada finalmente por uma equipe de mais de 45 cientistas dirigidos pela marca H. Skolnick do geneticist da universidade de Utá e de Miríade Genetics Inc. da cidade do lago salt. Uma outra equipe também identificou a posição aproximada de um segundo gene, BRCA2, que é responsável acreditado para a maioria das caixas restantes do cancer de peito herdado nas mulheres, possivelmente muitos casos esporádicos da doença, as well as os cancers de peito masculinos.

A descoberta de BRCA1 é " um avanço principal (que) permitir-nos-á de identificar mulheres, particularmente umas mulheres mais novas, que tenham um risco aumentado de desenvolver o cancer de peito, " biólogo molecular dito Harold Varmus, diretor dos institutos nacionais da saúde. Talvez mesmo mais importante, entretanto, seja a compreensão que nova os genes fornecem sobre como o cancer de peito ocorre, de acordo com o Dr. Harmon Eyre, oficial médico principal da sociedade americana do cancer. " nós sabemos realmente muito pouco sobre o que causa o cancer de peito, " Eyre dito. " esta é uma primeira etapa real em começar ao fundo de o que é toda sobre.

" As activistas do cancer de peito foram deleitadas pela descoberta, particularmente por causa de suas implicações potenciais para compreender a biologia da doença, disseram Michelle Rakoff, co-cadeira do alliance do cancer de peito de Los Angeles. " o que nós temos agora é a deteção adiantada, " disse. " o que nós queremos realmente é a prevenção, e para aquele nós necessitamos saber o que o causa.

" Os relatórios na descoberta foram programados para ser publicados mês seguinte na ciência do jornal, mas foram liberados quarta-feira depois que a notícia do NBC revelou seus índices. O cancer de peito é o formulário o mais comum do cancer nas mulheres, nas mulheres 182.000 estimadas golpeando de uns ESTADOS UNIDOS e nos 1.000 homens cada ano.

Estima-se que um em cada nove mulheres desenvolverá a doença sometime em suas vidas. A taxa da cinco-ano-sobrevivência depois que o tratamento para o cancer de peito localizado é 93%, mas a taxa mergulha a 72% se a doença espalhar na região do peito na altura do diagnóstico e a 18% se a doença espalhar a outras partes do corpo.

Da doença as matanças anualmente aproximadamente 46.000 mulheres e 300 homens. Os investigadores souberam por mais de 100 anos que o cancer de peito funciona nas famílias, com este formulário da contabilidade da doença para aproximadamente 5% de todos os casos. O formulário herdado é fisicamente indistinguishable do cancer de peito esporádico, mas dele greves em uma idade mais adiantada, com uma proporção grande dos casos que ocorrem nas mulheres sob a idade de 30.

Os estudos precedentes de mais de 200 famílias com o gene BRCA1 revelaram que o gene está ligado não somente ao peito e ao cancer ovarian, mas talvez ao cancer dos dois pontos e do prostate também. as mulheres e os homens nestas famílias têm uma taxa acima do normal do

cancer dos dois pontos e os homens têm uma quantidade excessiva de cancer do prostate. O gene recentemente identificado, Skolnick dito, é um gene muito grande " que não tenha nenhuma similaridade a nenhum outro gene que nós soubermos até do último segmento que nós encontramos." A estrutura desse segmento sugere que a proteína produziu dos ligamentos do gene ao DNA dentro da pilha, mas os investigadores ainda não tem nenhuma idéia o que faz depois que liga. A maioria de investigadores, entretanto, acreditam que BRCA1 é um gene do supressor do tumor -- um que mantém normalmente o replication das pilhas na verificação.

O grupo de Skolnick está trabalhando fora do tempo estipulado para localizar sua função quando saudável. " está indo ser difícil, para dizer o menos, compreender como este gene predispõe algumas mulheres ao peito e ao cancer ovarian, " o rei dito. " mas no one peito sempre dito ou cancer ovarian seriam fácil de resolver." Skolnick disse que requereria ao menos um ano para desenvolver um teste para o gene defeituoso porque os investigadores necessitam ser certos que os mutations causam a cancer, que são harmless e que são mais prováveis causar tumors ovarian do cancer melhor que do peito. " nós queremos certificar-se de que se você disser uma mulher tem um risco elevated, é verdadeiro, " disse. Baseado no tamanho do gene, um teste para detectar sua presença poderia custar tanto quanto \$1.000, ele disseram.

Eyre anotou que o teste fará também possível determinar quando o gene defeituoso não está atual. Em toda a família com muitas mulheres, disse ele, há uma possibilidade significativa que diversos desenvolverão o cancer de peito mesmo que não carreguem o gene. Em tais famílias, o teste deve reduzir a ansiedade mostrando que os membros não afetados não têm nenhum risco incomum para o cancer de peito, ele disse. Atualmente, algumas mulheres nas famílias com um history da doença submetem-se a mastectomies prophylactic para reduzir seu risco. Se o gene não estivesse atual, tais mastectomies seriam claramente desnecessários.

O grupo de Skolnick também procurou o gene em diversos pacientes com formulários esporádicos da doença e, a todos surpresa, era incapaz de encontrá-la. " que é completamente inesperado, " Eyre disse. Geralmente, disse ele, um gene que jogue um papel principal em um formulário herdado de uma doença é também importante em casos esporádicos. Os investigadores têm umas esperanças mais elevadas que o outro gene sob o estudo, BRCA2, estará ligado aos casos esporádicos uma vez que foi isolado. A posição aproximada de BRCA2 no chromosome 13 foi determinada por uma equipe dirigida pelo biólogo molecular Douglas F. Easton do instituto da pesquisa de cancer em Londres, que estudou 15 famílias do cancer de peito que não têm BRCA1. " nós pensamos que BRCA1 explica talvez a metade dos casos herdados e BRCA2 explica uma fração boa do descanso, " do ele dissemos. " mas nós pensamos que lá é ainda um outro gene a ser encontrado." Easton e seus colegas -- uma equipe que inclui o grupo de Skolnick -- estão usando agora as mesmas técnicas usadas em Utá isolar BRCA2 entre dos aproximadamente 65 genes situados no segmento pequeno do chromosome 13 que identificou.

* INOVAÇÃO: Traçar do gene está mudando a biologia, está cultivando-a e medicina. D1

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

5.

(86327)<DOCNO> LA091694-0142

OS CIENTISTAS DESTRAVAM O GENE ' FANTÁSTICO ' PARA IMPULSIONAR O CRESCIMENTO DE PLANTA

Os cientistas destravaram um gene que um dia poderia trazer umas orelhas de milho mais grandes, de mais rosas em cada arbusto e mesmo de florestas mais faster-growing, tudo controlando um hormone dentro de cada planta. " quando eu penso dos usos -- é fantástica, " disse Jędrzej

Szerszen, um pathologist de planta da universidade de estado de Michigan que ajudasse descobrir como manipular geneticamente o hormone natural do crescimento das plantas.

Em um estudo publicado hoje na ciência do jornal, Szerszen e os colegas Robert Bandurski e Krzysztof Szczygłowski isolaram e clonaram um gene que controlasse o hormone. "é emocionante," disse o Dr. Jerry Cohen do departamento de ESTADOS UNIDOS de agricultura, que a engenharia genética dita pôde começar a substituir produtos químicos do crescimento em dois a cinco anos. "esta é a primeira etapa." O crescimento de planta depende de um hormone chamado IAA.

Um gene chamado iaglu produz um enzyme que controle quanto IAA está livre dentro das plantas impulsionar o crescimento e quanto é Bandurski inativo, explicado, que pesquisou IAA por 30 anos. Os cientistas do estado de Michigan arrancaram o gene do iaglu fora do milho, clonaram e manipularam-no para produzir mais do enzyme vital. Injetaram então o tabaco, que cresce rapidamente e aceita a engenharia genética bem, com o gene. Em vez da planta normalmente unweildy, cresceram o tabaco mais bushier com mais filiais -- essencialmente começar a mesma quantidade em um espaço menor.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

6.

(41668)<DOCNO> LA031794-0242

Ò GENE DO CANCER DOS DOIS PONTOS ENCONTRADO; TESTES ESPERADOS LOGO

Com a descoberta de um segundo gene do cancer dos dois pontos, os investigadores relataram quarta-feira que têm identificado agora as causas de mais do que 90% do formulário herdado da doença. Entre eles, os dois genes são responsáveis para um em cada seis das 156.000 caixas novas do cancer dos dois pontos diagnosticadas cada ano. Parecem também esclarecer o tanto como como 30% de caixas (non-herdadas) esporádicas do cancer dos dois pontos.

Os investigadores esperam dentro de alguns meses desenvolver os testes de diagnóstico que mostrarão se um indivíduo tem um ou outro gene. Se um dos genes estiver atual, os doutores podem monitorar para tumores frequentemente, permitindo sua detecção quando forem ainda curáveis pela cirurgia. "nós podemos reduzir mortes do cancer no excesso 90% destas famílias perto," disse o Dr. Bert Vogelstein da universidade de Johns Hopkins, um co-líder de um dos dois grupos que relatam a descoberta hoje na natureza e na sexta-feira britânicas do jornal na ciência.

A descoberta poderia também conduzir às drogas anti-cancer novas dentro de três a cinco anos, microbiologist predito Richard Fishel da universidade de Vermont, um líder da outra equipe, que relatou seus findings na natureza. "eu sou muito confiável que na ordem curta nós poderemos desenvolver o therapeutics apropriado baseado em nosso conhecimento destes genes," ele disse. As mesmas duas equipes relataram a descoberta do primeiro gene do cancer dos dois pontos em dezembro. Ambos os genes, que são encontrados em chromosomes diferentes, ato como a função do spell-check em um programa processar de palavra, verificando recentemente synthesized o DNA para assegurar-se de que nenhum erro -- mutations -- ocorresse durante o proliferation celular.

Quando um ou outro gene é defeituoso, "você acumula estas alterações em uma taxa extremamente elevada e o cancer é o resultado," disse Fishel. Greves herdadas do cancer dos dois pontos geralmente antes da idade de 50. "este é um triunfo grande para a ciência," disse o departamento da saúde e da secretária Donna Shalala dos serviços do ser humano. "estas descobertas. conduzirá aos testes de seleção para indivíduos high-risk logo. Os doutores poderão conservar vidas incontáveis e impedir o sofrimento muito needless."

Shalala levantou também o interesse, compartilhado pelos investigadores, que a seleção para os genes poderia agudamente danificar a habilidade dos portadores -- povos que têm os genes -- de compra o seguro de saúde. Porque virtualmente todos que tem os genes desenvolverá o cancer, os portadores puderam ser excluídos da cobertura. Aproximadamente um em cada 200 povos tem um dos genes defeituosos, Vogelstein dito, fazendo lhe o defeito genetic o mais comum no mundo. Ambos os grupos disseram que identificaram tentativamente um terceiro gene que explicasse a maioria do 5% restante a 10% do cancer herdado dos dois pontos, chamado formalmente cancer hereditary dos dois pontos do nonpolyposis. Em 1991, outros cientistas anunciaram a descoberta de um gene que causasse um formulário mais raro do cancer herdado dos dois pontos, chamada o polyposis adenomatous familiar, que explica talvez 1% de todas as caixas do cancer dos dois pontos.

Os investigadores sabem bastante sobre como os dois genes trabalham porque Fishel e o colega Richard Kolodner, um geneticist no instituto do cancer de Dana-Farber em Boston, estudaram genes virtualmente idênticos nas bactérias e no fermento por quase 20 anos. Nas bactérias, os genes produzem duas proteínas que realizam a função do " spell-check ". A primeira proteína faz a varredura do DNA recentemente synthesized para verificar para ver se há todos os erros que ocorrerem durante sua síntese. Se um erro for encontrado, a segunda proteína liga ao primeira e repara o erro. Se o primeiro gene for defeituoso, os erros na formação do DNA não estão reconhecidos. Se o segundo gene for defeituoso, os erros não estão reparados. Uma ou outra maneira, erros acumula até que uma pilha do cancer esteja dada forma.

Os investigadores são confiáveis um processo idêntico ocorrem com os genes humanos, mas aquele não foi provado. A descoberta rápida dos genes humanos, Fishel dito em uma entrevista, veio aproximadamente por causa do conhecimento extensivo que tinha sido compilado já sobre os genes bacterianos, um exemplo clássico da pesquisa básica ao contrário de aplicado. " esta descoberta é um coup para o processo da pesquisa básica nos estados unidos, " Dr. notável George Vande Woude do instituto nacional do cancer.

Ambos os grupos estão desenvolvendo ativamente os testes de laboratório que poderiam ser usados selecionar membros de famílias afetadas. Que o processo é complicado, Fishel disse, porque ao menos os mutations uma dúzia diferentes do gene novo foram identificados e cada um tem que ser feito exame no cliente em toda a tela potencial. O teste espera-se custar inicialmente aproximadamente \$1.000, mais ou menos como o colonoscopy usado agora à busca para tumores dos dois pontos.

Tal testar será o mais apropriado nas famílias onde ao menos três membros foram diagnosticados com dois pontos ou os tumors rectal dentro de duas gerações. Fishel disse que testar deve ser conduzido " assim que possível, " preferivelmente logo após o nascimento, porque o greve da lata do cancer nos teens. Os portadores podem também ajustar seu lifestyle -- adotando um ponto baixo da dieta em gordo e em elevado na fibra, para o exemplo -- para reduzir riscos.

O custo da seleção espera-se declinar com o uso aumentado, permitindo eventualmente o teste ser aplicado à população inteira. Há alguma evidência que os genes podem jogar um papel em outros cancers também, including aqueles dos pulmões, do intestine, do uterus e dos ovaries. Tais tumors ocorrem em uma taxa elevada nas famílias com um history do cancer dos dois pontos, e as versões defeituosas dos genes foram identificadas em alguns destes tumors dos indivíduos. Os portadores do gene seriam monitorados também para o desenvolvimento destes tipos de tumors.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

7.

(82660)<DOCNO> LA090194-0279

GENE DEFEITUOSO DO ACHADO DOS CIENTISTAS LIGADO PARA DESCASCAR O CANCER

Os cientistas identificaram um gene defeituoso que parecesse causar uma tendência herdada ao melanoma mortal do cancer da pele e pudesse também jogar um papel no melanoma non-herdado. Em seu estado normal, o gene age como um freio no cancer. Mas povos que herdaram uma versão defeituosa perdem aparentemente a parte da proteção, fazendo os raramente vulneráveis ao melanoma, investigadores ditos. Aproximadamente 32.000 casos do melanoma e quase 7.000 mortes do melanoma esperam-se nos estados unidos este ano. Somente sobre 10% do melanoma ocorre nos povos com uma tendência herdada e é unclear que porcentagem de casos herdados é por causa do gene. As versões defeituosas do gene podem também ser envolvidas em muitos, ou nivele talvez a maioria, casos do melanoma non-herdado, investigadores ditos.

Naqueles casos, o gene seria herdado no formulário normal mas danificado mais tarde pela luz solar ou por outra causas. Estudar o gene pode eventualmente conduzir para melhorar tratamentos para a doença non-herdada, investigadores ditos. Mais imediatamente, a descoberta de um gene do susceptibility podia ser usada selecionar para povos no risco para o melanoma. Poderiam ser aconselhados fazer exame de etapas tais como limitar a exposição ao sol, mantendo-se a par possivelmente de toupeiras precancerous e usando-se sunscreen, os investigadores ditos.

O gene normal diz ao corpo como fazer uma proteína chamada p16, que as ajudas regulam a divisão da pilha. Os estudos prévios sugeriram que o gene p16 desanima tumors. Os estudos precedentes indicaram também o jogo defeituoso das versões um papel no cancer. O trabalho novo é relatado na introdução de setembro do genetics da natureza do jornal por duas equipes independentes dos investigadores. Um estudo foi feito por Nicholas Dracopoli do centro nacional para a pesquisa humana do genome, a parte dos institutos nacionais da saúde em Bethesda, o Md., e os colegas em NIH e em outra parte.

Encontraram nove de 18 famílias que melanoma-melanoma-prone selecionaram defeitos mostrados no gene p16 que pareceu relacionado à doença. Nas nove famílias, o gene era defeituoso em 33 fora de 36 povos com melanoma. O segundo estudo foi feito por Alexander Kamb de Miriade Genetics Inc. na cidade do lago salt com os cientistas na universidade de Utá e em outra parte. Estudaram 13 famílias que pareceram já ter um predisposition ao melanoma de algum abnormality na vizinhança do gene p16. Os investigadores encontraram um gene p16 defeituoso em duas das famílias. Não encontraram nenhuma evidência de um gene p16 defeituoso outras em 38 famílias melanoma-melanoma-prone. Nonetheless, o Utá e o NIH estudam " realmente a sustentação, fundamental, " Kamb dito, anotando que dois tipos de defeitos do gene p16 apareceram em ambos os estudos.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

8.

(109939)<DOCNO> LA121894-0057

MEDO NOVO DA CIÊNCIA DE ALIMENTO: ALLERGEN DE STEALTH; BIOLOGIA: A ENGENHARIA GENETIC FARÁ MAIS DIFÍCIL DE SABER O QUE É REALMENTE DENTRO DO PRATO DO CASSEROLE.

Se você for allergic às porcas, é fácil bastante passar acima da mistura do cocktail. Mas não pode ser assim que fácil para que os povos passem acima de algumas das misturas allergy-allergy-triggering do potencial da manipulação genetic do milho, dos tomates, do squash e de incontável

outros alimentos. Que se, para o exemplo, algo de uma porca de Brasil for posto em alguns soybeans, que terminam acima em tudo da fórmula do bebê à Bolonha?

Que o tipo da pergunta pareceria foolish antes que os cientistas aprenderem mover o material genetic da planta para a planta, animal à planta, bactérias à planta, vírus a plantar. Os reguladores do governo, as companhias do biotech e os médicos cresceram cada vez mais interessados sobre o prospeito de baralhar allergens na plataforma do DNA de um alimento. Apesar de tudo, o preço de noshing wrongful funciona dos hives a um ataque life-threatening do asma para alguns dos milhões dos americanos que são allergic às porcas, aos ovos, aos amendoins do leite e aos outros alimentos. " qualquer coisa que você transfere aumentos uma pergunta, " disse o Dr. Hugh A. Sampson, specialist do allergy do pediatrician e do alimento na escola da universidade de Johns Hopkins da medicina. " você começou apenas decidir-se como é provavelmente a ser um problema."

Os cientistas esperam que adicionando, suprimindo ou mudando traços hereditary possam desenvolver as colheitas que provam mais melhor, permanecem mais longo fresco, oferecem mais nutrientes e poucas gorduras, resistem pests, ou toleram herbicidas. Mas o esforço pela companhia a maior da semente do mundo embalar mais nourishment em um soybean mostra o que uma engenharia genetic do negócio risky pode ser -- e como bom poucas proteções trabalharam assim distante. Os investigadores em Pioneiro Hi-Bred Internacional Inc. de DES Moines souberam que as porcas de Brasil têm o methionine, um amino-ácido enxôfre-rico que os animais domésticos e as aves domésticas necessitassem. Os soybeans são uma alimentação dos animais domésticos do top, mas são pobres no enxôfre.

Adicionar o methionine faria os soybeans que mais de a superfeed. Mas o pioneiro soube sobre o problema do allergy. O alimento e a administração da droga disseram companhias do biotech em maio, 1992, para verificar para ver se há allergens quando testa a segurança de seus produtos. O pioneiro pediu ajuda do alfaiate de Stephen, cabeça do departamento da ciência de alimento da universidade de Nebraska em Lincoln. O alfaiate tentou o produto no serum do sangue dos povos com allergies da porca e encontrado havia uma reação. " nós estamos avaliando o projeto de pesquisa para o mais melhor trajeto para a frente, " disse Steve Daugherty, spokesman da companhia.

As opções incluem encontrar uma maneira transferir a proteína sem o allergen, ou deixar cair o projeto completamente. O teste mostrou-o que é possível selecionar alimentos bio-projetados para allergens sabidos. Prende também para fora da promessa que os cientistas podem usar o biotechnology começar snipping para fora dos allergens. Mas o biotechnology levanta também o specter de criar ou de adicionar substâncias allergy-allergy-inducing novas à fonte de alimento das fontes como as proteínas viral do revestimento que podem ajudar a uma planta resistir a doença.

Muitos cientistas discutem que os riscos são baixos, particularmente desde que parece que somente um punhado dos dez dos milhares das proteínas para fora de lá causa allergies. Os materiais novos introduzidos pela engenharia genetic são adicionados em níveis baixos e não podem estar atuais na parte edible do alimento. Ainda, o ignorance sobre allergies do alimento faz difícil de começar tão claramente um retrato dos perigos como a maioria de cientistas gostariam. " não há nenhum teste foolproof " para allergens suspeitados, disse o alfaiate, que nonetheless acredita que os riscos são aceitáveis. Os cientistas estimam que os allergies do alimento afflict sobre 2% de crianças novas e de 1% dos adultos.

Top dos amendoins, do leite, dos ovos, dos soybeans, dos peixes e do marisco a lista de alimentos allergy-allergy-causing neste país. A volta do biotechnology podia ser um impulso para investigadores do allergy do alimento. " porque você sairia e gastaria milhões dos dólares que isolam dúzias de allergens do alimento quando você não tiver nenhuma finalidade para ela?" Dr. dito decano Metcalfe, do instituto nacional do governo do allergy e de doenças infectious. " de repente, agora, você necessita-o."

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

9.

(112520)<DOCNO> LA123094-0055

CIÊNCIA / PESQUISA GENETIC; O ESTUDO LIGA O ABNORMALITY A ATROPHY MUSCULAR SPINAL

Os investigadores UC de Irvine relatam que identificaram um abnormality que genetic acreditam causas atrophy muscular spinal, a causa genetic a mais comum da morte nos infantes. " este é encontrar longo-antecipado que deve rapidamente melhorar a exatidão no diagnóstico dos formulários de SMA, " disse Donald S. Madeira, diretor da tecnologia da ciência para o dystrophy muscular Assn. " o mais importante, acelera nossa busca continuando para um formulário do tratamento para SMA." Nenhum tratamento existe agora. O problema, que em seu formulário mais severo é uniformemente fatal antes da idade de 2, afeta os 20.000 americanos estimados, fazendo lhe o problema genetic letal o mais comum do segundo após o fibrosis cystic. Afeta um em 10.000 nascimentos, mas estimado em 40 adultos carrega o gene defeituoso. A identificação do gene é relatada por uma equipe conduzida pelo biólogo molecular John J. Wasmuth de UC Irvine na introdução de janeiro do genetics da natureza do jornal, liberada hoje. A equipe de Wasmuth acredita que o problema está causado pela perda dos segmentos de um gene chamado XS2G3 no chromosome 5, um dos 23 pares dos chromosomes que fazem acima o blueprint genetic humano. Estudaram o DNA de 29 pacientes com o formulário o mais severo da doença, chamaram Tipo Eu SMA, e compararam-no ao DNA de 235 povos saudáveis. Encontraram que 17 dos 29 pacientes eram fragmentos faltantes de XS2G3 em ambas as cópias do gene (um de cada pai). No contraste, somente dois dos 235 povos saudáveis -- ambos eles portadores sabidos do problema -- eram fragmentos faltantes do gene, e somente em uma cópia. A madeira muscular do dystrophy Assn.'s speculated que os fragmentos faltam em todos os 29 pacientes mas que os investigadores eram incapazes de os detectar em 12 casos por causa da complexidade extrema dessa área do chromosome 5.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

10.

(88274)<DOCNO> LA092394-0150

OS PRESERVATIVOS DE ALIMENTO ENCONTRARAM PARA IMPULSIONAR LUTADORES DO CANCER

Dois níveis impulsioneados extensamente usados dos preservativos de alimento de um lutador natural do cancer em animais de laboratório e parecem fazer a mesma coisa nos seres humanos, um investigador relatado quinta-feira. Os advogados de alimentos naturais decried por muito tempo o uso dos preservativos, mas o Dr. Andrew Dannenberg da faculdade médica de Cornell encontrou que os preservativos BHA e BHT " revved acima " do gene para um enzyme que as ajudas destruísem carcinogens antes que provoquem tumores.

Quando os genes são postos em marcha acima, produzem mais do enzyme, fornecendo a proteção melhor de encontro às substâncias cancer-cancer-causing no ambiente, Dannenberg relatado na conferência internacional sobre a prevenção do cancer na universidade de Rockefeller em york novo. O BHA e o BHT são aditivos sintéticos usados como preservativos nos bolinhos, nos biscoitos e em uma variedade larga de outros alimentos.

Os resultados não significam que os alimentos devem ser bombeados acima com mesmo mais preservativos, Dannenberg dito. Os findings são importantes porque descobrem um mecanismo da prevenção do cancer que pareça ser parte da explanação para as propriedades anti-cancer well-

known dos bróculos, da couve-flor e dos sprouts de Bruxelas. Sua pesquisa mostra que ao menos a parte do efeito daqueles compostos vem de revving acima do mesmo gene afetado por BHA e por BHT.

O gene produz um enzyme com o nome longo UDP-glucuronosyltransferase, ou UGT. O estudo encontrou níveis elevated do enzyme no figado, kidneys e os intestines pequenos dos ratos alimentaram uns doses mais elevados do BHA e do BHT do que são encontrados normalmente nos alimentos, Dannenberg dito. Encontrou então a evidência preliminar que as substâncias fazem a mesma coisa nos seres humanos. Dannenberg disse que encontrou também esse sulforaphane, um agente anti-cancer isolado recentemente nos bróculos, exerce sua ação em parte na mesma maneira, revving acima do gene para UGT.

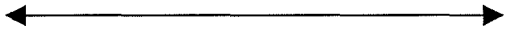
Os resultados aparecem em uma hora quando os investigadores se tornaram convencidos extensamente do valor das frutas e dos vegetais no cancer impedindo mas fizeram pouco headway em figurar para fora exatamente porque aquele é verdadeiro. Um problema é que muitos povos evitam os alimentos com os efeitos anti-cancer os mais potent, disse o Dr. Walter Willett da escola de Harvard da saúde pública.

Em relação ao tópico, este documento é:

- () relevante
- () irrelevante
- () não sei


Em uma escala de 1 a 5, como você classifica a dificuldade da tarefa para este tópico (onde 1 significa “muito fácil” e 5 significa “muito difícil”)?

Muito Fácil 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Muito Difícil



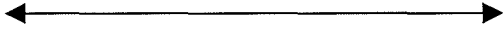
Em uma escala de 1 a 5, o quão confiante você está quanto à precisão dos julgamentos que fez para este tópico (onde 1 significa “muito confiante” e 5 significa “muito inseguro”) ?

Muito Confiante 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Muito Inseguro



Como você considera a sua familiaridade com o tópico da consulta (onde 1 significa “especialista neste assunto” e 5 significa “não tenho nenhum conhecimento do assunto”)?

Especialista 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Nenhum conhecimento



Casamento de Jackson e Presley

Encontrar documentos sobre o presumível casamento de Michael Jackson com Lisa Marie Presley, ou sobre a sua separação.

Em Maio de 1994, a famosa estrela da música pop Michael Jackson, foi dada como tendo casado com Lisa Marie Presley, a filha do rei do rock-and-roll. Documentos relevantes devem conter detalhes sobre o casamento, tal como por exemplo onde ele terá sido celebrado, ou devem discutir sobre a separação posterior do casal.

1.

(31172) <DOCNO> LA020694-0059

CREMALHEIRA RECORD: NO SUMÁRIO

** Freddie Jackson, " aqui está, " Jackson de RCA. Poder seja o Michael Bolton da música preta? Para trás de um hiatus longo da gravação, de Jackson dos oversings realmente ao índice do seu coração em diversos cuts."Was algo " e " como sente o " fare bem sob tal tratamento, mas de his que overemoting no ballad amor-v-errado " minha família " não faz o sentido. Como Bolton, Jackson falta muito no departamento do sincerity. Os albums novos são rated em uma escala de uma estrela (pobre), de duas estrelas (justas), de três estrelas (boas) e de quatro estrelas (excelentes). JOHNSON DE CONNIE

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

2.

(49220) <DOCNO> LA041794-0218

MICHAEL JACKSON COMEÇA ' IMPORTANDO-SE COM A CONCESSÃO DOS MIÚDOS

Michael Jackson pode ter tido sua parte de problemas legais sobre allegations do abuso da criança, mas 100.000 estudantes da área de york novo deram-lhe um voto da confiança, selecionando o receber " importando-se com uma concessão dos miúdos ".

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

3.

(75853) <DOCNO> LA080394-0051

AS FAÍSCAS DA UNIÃO DE JACKSON-PRESLEY CHOCAM-SE, DUVIDAM-SE, RISOS; UNIÃO: CONFIRMAÇÃO QUE O REI ESTALAR E PRINCESA ROCHA TER CERTAMENTE ENGATOU CAUSAS MUITAS REAÇÕES, COM TOPPING DA SURPRESA

A CARTA. Os ventiladores de Elvis Presley e de Michael Jackson foram unidos em uma coisa terça-feira: seu choque sobre a confirmação da união do rei do PNF à filha do rei da rocha. "ninguém pensou sempre que veriam o dia quando Michael Jackson viria moonwalking abaixo o bulevar de Elvis Presley, " disseram Janis Fullilove, um de rádio fal-mostram o anfitrião em WMC no hometown de Elvis' de Memphis.

No hometown de Jackson de Gary, o Ind., a conta James, que hospeda uma mostra da conversa em WLTH, relatou que a reação à união de Jackson e de Lisa Marie Presley era skeptical. de " os ventiladores Jackson não acreditam que está casado, " James dito. " pensam que é justo para o publicity." O trabalho de Bob, presidente do estado Assn. de 367-member Jersey nova para Elvis, disse os poses da união um quandary.

Quer saber como está indo segurar a história em tempos de Elvis, o boletim de notícias que editou cada duas semanas por 12 anos. " eu não estou indo tocar ainda n, " trabalho dito hesitantly pelo telefone terça-feira. " é demasiado touchy. Você tem os ventiladores que são afeiçãoados de Lisa e irão junto com qualquer coisa que , e os outros que não são demasiado afeiçãoados do publicity de Michael, e o problema está sendo começado com em miúdos. " todos quiet, " trabalho disse da sociedade do clube de ventilador. " estão em choque.

Naturalmente, a outra coisa em todos mente é, onde Graceland caberá na maneira traga a estrada? O que a vontade Michael quer fazer com ela?" " Cynthia Horner, diretor editorial para a direita teen preta do compartimento sobre!, não pensa de Lisa casado Michael Marie apenas para começar suas luvas em Graceland, ou nivele para melhorar sua imagem pública negativa. " Michael pode começar o publicity quando quer, " disse Horner, que soube a estrela do PNF desde 1976. " não é necessário que faça exame desta rota à tentativa e à carne acima de sua reputação.

É demasiado confidencial de uma pessoa. " eu falei a Michael em abril, e nesse tempo indicou procurava uma esposa, e eu lhe disse que o procurar alguém ingualmente tão famoso e rico como é, assim que não teria que se preocupar sobre alguém que casa o apenas para seu dinheiro." Em cima do hearing seu conselho, Horner relatou que Jackson " sorriu apenas." Em registros da torre na noite de Hollywood segunda-feira, a reação aos nuptials de Jackson-Presley era mais cínica. Como Calvin Pixley, 37, dito: " eu não v o amor lá. Começaram casados porque são ambos famosos?

Eu penso que admirou realmente a fama (de Jackson) -- seu pai que são assim que famoso e tudo -- e caiu no amor com sua arte, sua fama." " é absurdo, " disse Darrin Tehrani, 29, um músico que creditasse Jackson e Presley entre his influências. " uniforme se for verdadeiro, que são seus motriz? São dois oddballs e esta coisa inteira é um mess. Eu estou indo provavelmente ter sonhos maus sobre este." Na mostra do dirij-tempo da manhã de KROQ's, os anfitriões Kevin e o feijão gastaram teorias fielding do ouvinte de terça-feira em como e em porque o rei do PNF e da princesa da rocha começou junto. A idéia de Kevin do pensamento do feijão estava distante pelo mais estranho. " pensa realmente de que talvez se gostam de." Kevin Reconhecido: " está refrescando para não considerar que são justos fazendo a para o dinheiro. ao contrário de Tom e de Roseanne ou de Anna Nicole (Smith) e grandpa."

Em relação ao tópico, este documento é:

- () relevante
- () irrelevante
- () não sei

4.

(23290) <DOCNO> LA010594-0010

SUMÁRIO METROPOLITANO / NOTÍCIA CONDADO DE LOS ANGELES NO SUMÁRIO:
LOS ANGELES; O ADVOGADO DO MENINO PROCURA FOTOS DO CORPO DE
MICHAEL JACKSON

O advogado que representa um menino 13-year-old que o reivindicasse foi molestado por originais arquivados Jackson terça-feira da corte de Michael em um esforço obter fotografias do corpo do anfitrião. O último mês, Jackson submeteu-se a uma busca do corpo pelos investigators que procuram a evidência corroborate as reivindicações do menino. " nós pensamos de que o fato que meu cliente pode estabelecer que olhares do Sr. Jackson como despido são evidência muito substancial da culpa do Sr. Jackson, " disse Larry Feldman, advogado do menino.

Feldman disse que arquivou um movimento na corte que é um pedido " bem escolhido " múltiplo: Jackson pode fornecer cópias das fotografias das polícias, submete-se a uma segunda busca, ou a corte pode barrar as fotografias da experimentação civil como a evidência. Feldman disse pediu que advogados de Jackson e o escritório do advogado do distrito do condado de Los Angeles para cópias, mas elas recusou.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

5.

(35342) <DOCNO> LA022094-0362

MICHAEL APARECE MOMENTANEAMENTE ' NA MOSTRA DAS HONRAS DA FAMÍLIA
DE JACKSON; ENTERTAINMENT: A ESTRELA DO PNF APRESENTA CONCESSÕES À
BAGA GORDY DO ALFAIATE DE ELIZABETH E DO FOUNDER DOS REGISTROS DE
MOTOWN.

Faturado como um evento humanitário, de " as honras da família Jackson " que gravam aqui sábado tiveram mais da sensação de um melodrama da tevê. O membro o mais famoso da família, a estrela beleaguered Michael Jackson do PNF, parecido ser um participant relutante na produção mais do que two-hour, e no membro o mais mercurial da família, LaToya Jackson, contended que esteve barrada do programa. Em vez de executar um número de solo -- como a irmã Janet Jackson, que recebeu um ovation ereto cedo na noite -- Michael Jackson concordado somente às concessões atuais fechar o alfaiate de Elizabeth do amigo e os registros de Motown founder a бага Gordy.

Sua somente contribuição musical programada à mostra juntava a família em um finale escrito para o evento, que deve ser transmissão terça-feira por NBC-TV. Porque a foto do singer 35-year-old tinha sido caracterizada proeminente no ads para a mostra do reunion, muitas na multidão da próximo-capacidade na arena grande do jardim de MGM 13,000-seat tinham esperado para mais música dele. " a razão inteira I pagou \$100 por um bilhete devia ver Michael executar, " Julie Marlowe, 24, dito. " I pensou que este esteve suposto para ser um concert." Ainda, esta era a maioria de aparência do estágio de ESTADOS UNIDOS do elevado-perfil por Jackson desde que foi acusado do molestation sexual por um menino 13-year-old em agosto, e a maioria da multidão entusiástica estava ansiosa para mostrar sua sustentação. " eu tenho amei sempre sua música, e eu quis mostrá-lo que nós nos importamos ainda com ele, " disse Maria Lopez, 26, após ter comprado uma tarde de sábado de \$100 bilhetes.

Jackson pagou reportedly a mais de \$10 milhões o último mês para estabelecer o lawsuit. Nenhuma carga do criminoso foi alojada de encontro a ele. Durante jogar de um de seus videos para abrir a mostra de 8 p.m., a multidão shrieked na aprovação. O vídeo mostrou que Jackson que executa o tune " sae de me sozinho, " qual foi significado mock seu tratamento pela imprensa do tabloid. Backstage antes de apresentar sua primeira concessão, Jackson expressou a gratitude para a lealdade dos seus ventiladores. " alguns amigos são como sombras, você vêem-nas somente quando o sol brilha, " que ele disse. " mas meus ventiladores sustentaram-me mesmo nos dias escuros. Eu devo-lhes tudo." Em uma nota mais menos harmonious, LaToya Jackson mantido mais cedo na semana que a família não a deixaria juntar a mostra a menos que assinar uma ordem da mordaga que concorda não criticar mais mais Michael Jackson ou família.

Em uma conferência da notícia da manhã, Jermaine Jackson, a força guiar atrás da tevê especial, e mãe, Katherine, negou que uma ordem da mordaga tinha sido exigida. Embora houvesse os boatos que LaToya Jackson pôde tentar sneak no concert no disfarce, seu marido, Jack Gordon, disse sábado que sua esposa tinha rejeitado a idéia depois que a sugeriu.

Apesar de todo o foco na família de Jackson, o evento de sábado não foi projetado como um reunion de Jackson, Jermaine Jackson dito. Instead, foi planeado como um evento anual saudar humanitários proeminentes e levantar o dinheiro para charities. Aproximadamente \$500.000 earmarked para o charity, na maior parte para o relevo do terremoto em Califórnia, disseram representantes da família de Jackson.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

6.

(107793) <DOCNO> LA121094-0112

FOCO DE VALLEYWIDE: A FAMÍLIA DE JACKSON SUED POR APPAREL FABRICANTE

Acusando a família de Jackson de transações " duplicitous " do negócio, os advogados para um fabricante francês da roupa arquivaram um lawsuit que procura mais de \$17 milhões, contending que os anfitriões reneged em um negócio exclusivo para introduzir no mercado o nome e o logo de família de Jackson em uma linha das calças de brim e do sportswear.

No terno, a companhia da roupa de Nelson de Toulouse, França, alega que a família de Jackson violated um junho, um 1991, um contrato Nelson concedendo " uso exclusivo do nome e um logo ' da família de Jackson ' em alguns e em todos os produtos manufaturados e distribuiu-o " por Nelson. Nomeado no terno que alega o fraud e a ruptura de contrato eram patriarch Joseph Jackson da família de Jackson Comunicações Inc., superstars Michael e Janet Jackson, as well as outros membros da família including Tito, Marlon, LaToya e Germaine Jackson.

A escova de Michael dos advogados e Michael que os sacks contend que após o Jacksons assinou o contrato que dá direitas exclusivas a Nelson e a seu presidente, Tony Elichia, família deram forma a uma companhia suíça chamada grupo unido dos talents e começaram a soliciting outro, arranjos similares do negócio. Os advogados disseram que o Jacksons era \$100.000 pagos para as direitas licenciando e os pagamentos adicionais do royalty que atingiram centenas dos milhares dos dólares. Muitos daqueles pagamentos estavam no dinheiro, escova dita.

Os representantes da família de Jackson não podiam ser alcançados. Entretanto, Jay Lavelly, um advogado para Michael Jackson, disse que seu cliente " lhe nega fêz qualquer coisa wrongful, fraudulent ou impróprio " e esse Michael Jackson nunca concedeu à companhia de Nelson a direita usar seu nome ou likeness para finalidades comerciais.

De acordo com uma exibição da corte unida ao lawsuit, Joseph Jackson escreveu uma letra em março, 1994, terminando o acordo. O Jacksons foi batido durante o ano passado com diversos lawsuits que alegam a ruptura de contrato em relação à televisão financeira incomodada das honras da família de Jackson especial. Além, Michael Jackson pagou a milhões mais cedo este ano para estabelecer um lawsuit arquivado por um menino 14-year-old que acusasse Jackson sexually de o molestar. ANN O'NEILL

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

7.

(82283) <DOCNO> LA083094-0190

NOTÍCIA DE METRO: ALTURAS DE LADERA; A ESPOSA ANTERIOR DE TITO JACKSON AFOGA-SE NO POOL

Delores Jackson, ex-esposa do irmão Tito de Michael Jackson, afogada acidentalmente em um pool de natação ao visitar o repouso das alturas de Ladera de seu boyfriend, o departamento do sheriff do condado de Los Angeles relatou. Jackson, 39, tem nadado com seu boyfriend, Donald Bohana, 58, em seu repouso na avenida de Holt aproximadamente 4 a.m. Saturday. Bohana saiu do pool e à esquerda momentaneamente. Quando retornou, encontrou Jackson no fundo do pool, deputado Rico Erickson do sheriff dito.

Erickson disse Jackson puxado Bohana da água e chamou 911. Os deputados não souberam quanto tempo tinha estado na água. Jackson foi feito exame pela ambulância ao hospital memorial de Daniel Freeman em Inglewood, onde era mortos pronunciados. os " investigators acreditam Jackson morrido de afogar-se acidental, mas a causa exata da morte será determinada pelo examinador médico, " Erickson dito. Delores Jackson divorciou Tito Jackson em 1988 após 16 anos da união.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

8.

(82053) <DOCNO> LA082994-0164

EX-WIFE DE TITO JACKSON AFOGA-SE NO POOL DE NATAÇÃO

Delores Jackson, ex-esposa do irmão Tito de Michael Jackson, afogada acidentalmente em um pool de natação ao visitar o repouso das alturas de Ladera de seu boyfriend, o departamento do sheriff do condado de Los Angeles relatou domingo. Jackson, 39, tem nadado com seu boyfriend, Donald Bohana, 58, em seu repouso na avenida de Holt aproximadamente 4 a.m. Saturday. Bohana saiu do pool e à esquerda momentaneamente. Quando retornou, encontrou Jackson no fundo do pool, deputado Rico Erickson do sheriff dito. Erickson disse Jackson puxado Bohana da água e chamou 911.

Os deputados não souberam quanto tempo tinha estado na água. Jackson foi feito exame pela ambulância ao hospital memorial de Daniel Freeman em Inglewood, onde era mortos pronunciados. os " investigators acreditam Jackson morrido de afogar-se acidental, mas a causa exata da morte será determinada pelo examinador médico, " Erickson dito. Delores Jackson divorciou Tito Jackson em 1988 após 16 anos da união.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

9.

(83398) <DOCNO> LA090494-0058

OLHO DO PNF E UMA MAIS COISA:

" eu não tive uma moeda de um centavo (de minhas canções) desde 1980, " diz Neil Innes, em seu guise de Ron nasty, que fronteará " o Rutles novo " nas mostras quinta-feira e (adicionado recentemente) sábado no Troubadour. " agora parece que Michael Jackson comprou meu publicar por mistake. You não começou seu número de telefone, têm você?" Bem, o Rutles era um spoof do Beatles cometido pelo membro anterior da faixa do cão de Bonzo -- e sometime Monty Python co-conspirator-conspirator -- Innes com idle de Eric de Python.

Mas este comentário não é assim muito de um gracejo. Parece que isso devido às referências intencionais às canções de Beatles nos tunes que Innes escreveu para o Rutles, ele teve que trabalhar um negócio com ATV, então-proprietário da canção de Fab quatro publicar, para alguns dos royalties. Diversos anos há, naturalmente, Jackson comprou ATV, pela maior parte porque desejou ao próprio as canções de Beatles. E com ele comprou, se o soube ou não. o Rutles Canon. Um spokesman de Rutles diz que Jackson é bem-vindo nas mostras do grupo, e assim que é sua esposa, Lisa Marie Presley, para essa matéria.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

10.

(75852) <DOCNO> LA080394-0050

O QUE SE TRANSFORMA UMA LEGENDA MAIS: NOTORIETY, MÚSICA?; COMMENTARY: O HOOPLA SOBRE A UNIÃO DE JACKSON-PRESLEY PODIA MAIS MAIS OBSCURECER SEU FOCO CREATIVO. SOMENTE LATA LEGITIMIZE DA QUALIDADE MUSIC-MAKING SEU REI ESTALAR REIVINDICAR.

E nós pensamos que Julia e Lyle eram o par impar final ' do 90s. A confirmação desta semana de Michael Jackson e de casamento de maio de Lisa Marie Presley faz exame de nos além do reino do meramente quirky e no mundo do genuína bizarre. Aqui você tem Jackson, atualmente a estrela a mais famosa do PNF no mundo -- fresco do humiliation de cargas do molestation da criança -- na necessidade desesperada de re-establishing seus credibility e imagem.

Casa a filha da estrela a mais famosa do PNF de toda a hora -- em linha reta de seu próprio divórcio heartbreaking -- que quer saber talvez se puder sempre encontrar alguém que compreende as pressões de viver entre o scrutiny constante que vem com ser um Presley. Agora que o casamento esteve confirmado, a maioria de povos não podem começar bastantes da história -- mesmo que não tem nada fazer com porque nós nos importamos com Jackson no primeiro lugar: seu artistry.

O perigo para Jackson é que todo este fascínio público poderia mais mais obscurecer seu próprio foco criativo e destruir todo o credibility musical lingering. Em anos recentes, Jackson pareceu tão interessado em aparecer nos tabloids quanto nas cartas. A coisa engraçada é que a história inteira da união esteve colocada para fora das semanas há nos tabloids mas que ninguém fizeram exame d seriamente. A razão que a imprensa mainstream colocou fora o boato da união não era justa que os conselheiros de Jackson negaram toda à união.

Ninguém acredita celebridades ou seus conselheiros anymore de qualquer maneira. Esta era uma história que fosse simplesmente demasiado preposterous. O rei de começar do PNF casou-se? À filha do rei da rocha? Se a fizessem em um filme, teriam que pô-lo sobre o canal de Sci-Fi. O caso inteiro dá o meaning novo " às mentes suspicious batidas Elvis velhas." Entre as perguntas levantou sobre este merger unprecedented do PNF: É amor? É relações públicas? É terapia nova da idade? É um coup de Scientology? (Lisa Marie é um seguidor.) Graceland transformar-se-á agora um museu de Elvis Presley e de Michael Jackson? Se for difícil fazer o sentido fora das vidas das celebridades que dão entrevistas infinitas, contradictory (yo, Roseanne!), imagine o desafio de figurar para fora do motivation dos povos que não dão alguns.

Como você começa a remendar junto pensar atrás desta união? Uma vez os gracejos funcionados para fora, você começa a realizar por eles é dois dos povos os mais confidenciais e os mais famosos no mundo. Sua experiência no spotlight do PNF -- não obstante se Lisa Marie pisou sempre em um estágio -- dá-lhes uma ligação original. Para Lisa Marie, cujos birthright e a fortuna pode às vezes parecer tanto quanto um burden quanto um blessing, Michael poderia estar em seus olhos como alguém que gastou sua vida com o mesmo jogo dos riches e de tensões emocionais.

A parte da apelação original de Lisa Marie para Michael é seu bloodline. Foi um de poucos povos no mundo do PNF que poderia fazer Michael mais famoso do que antes que -- e você não dub o rei do PNF se você não obsessed com fama. Michael acredita que o one-way para realçar sua fama deve manter povos curiosos. Elvis Presley e Howard Hughes era ingualmente fascinante a ele porque eram originais. misterioso.

Uma razão Michael deu não falando à imprensa é, " Elvis não fez entrevistas." Mas há um perigo em tal ênfase na fama -- e o trabalho de Michael sofreu dele em anos recentes. Parece pensar dessa fama, melhor que a qualidade da música, é o que importa finalmente. Visto que havia um sentido do revelation e da paixão phenomenally no album bem sucedido do " thriller 1982's ", seus dois albums subseqüentes foram calculados mais comercialmente do que inspirados. Os albums ambos pareceram projetados proteger sua parte de mercado apelando aos vários formatos de rádio e estalar genres. Nenhuma venda da maravilha deslizou dramática nos estados unidos.

Demasiada da música pareceu artificial e fria. Era unreasonable pensar de que Jackson poderia sempre top os 20 milhão-mais vendas do " thriller " nos estados unidos, mas não era fora da linha para esperar que poderia nos dar um outro album com coração e fogo iguais. Que povos recordam sobre Elvis não é quanto o no. 1 o grava teve mas sua influência na música de PNF. Elvis não inventou o rolo da rocha ' n ', mas definiu-o nas maneiras que tocaram em todos do Beatles a U2. Similarmente, é a música que deve importar no exemplo de Jackson -- egardless do fascínio inevitável com sua vida pública.

Os povos perto de Jackson são ecstatic sobre as trilhas que novas fez para " o album das batidas as mais grandes " devido em torno Nov. da 1. Dizem que é alguma da música a mais forte desde o " thriller." Mas estes são alguns dos mesmos povos que negaram por semanas os boatos da união. Na extremidade, é a qualidade da música de Jackson que determinará se o título do rei do PNF é qualquer coisa mais do que um slogan vazio do marketing.

Appendix B – Relevance Feedback Questionnaire

Em relação ao tópico, este documento é:

- () relevante
() irrelevante
() não sei

Em uma escala de 1 a 5, como você classifica a dificuldade da tarefa para este tópico (onde 1 significa “muito fácil” e 5 significa “muito difícil”)?

Muito Fácil 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Muito Difícil



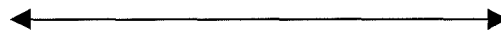
Em uma escala de 1 a 5, o quão confiante você está quanto à precisão dos julgamentos que fez para este tópico (onde 1 significa “muito confiante” e 5 significa “muito inseguro”) ?

Muito Confiante 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Muito Inseguro



Como você considera a sua familiaridade com o tópico da consulta (onde 1 significa “especialista neste assunto” e 5 significa “não tenho nenhum conhecimento do assunto”)?

Especialista 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Nenhum conhecimento



Morte do líder do Nirvana

Como morreu o cantor líder do grupo de rock e grunge americano Nirvana?

Kurt Cobain, cantor líder dos Nirvana, o famoso grupo de música popular, morreu em Abril de 1994. Documentos que relatem a morte do Cobain sem mencionarem as causas não são relevantes.

1. (64454)<DOCNO> LA061894-0109

KRISTEN PFAFF – DA BANDA DE COURTNEY LOVE, ENCONTRADA MORTA; MÚSICA POP: O CORPO DA BAIXISTA DO THE HOLE FOI ENCONTRADO NA BANHEIRA DE SUA CASA EM SEATTLE. A CAUSA DA MORTE NÃO FOI DETERMINADA

Kristen Pfaff, a baixista da aclamada banda de rock The Hole, foi encontrada morta em sua casa em Seattle na quinta-feira, de acordo com a Polícia de Seattle.

A banda The Hole é liderada por Courtney Love, a viúva do líder do Nirvana Kurt Cobain, que morreu após ter atirado contra si mesmo em Abril passado.

O corpo de Pfaff, 27, foi encontrado em sua banheiro por uma amiga. O relatório da Polícia conta que em uma bolsa próxima ao corpo havia seringas e outros objetos para consumo de drogas. Uma autópsia foi completada na sexta-feira, mas o porta-voz do investigador de homicídios do condado de King disse que a causa da morte não havia sido determinada, aguardando resultados de testes toxicológicos, os quais devem ficar prontos em 4 ou 6 semanas.

Um porta-voz da gravadora da banda, Geffen Records, disse que Pfaff havia completado um programa de reabilitação de drogas no último inverno e estava se preparando para mudar de volta para sua antiga casa em Minneapolis para esperar até que a banda The Hole recomeçasse a turnê que havia sido interrompida desde a morte de Cobain. Os membros sobreviventes da banda, cujo álbum "Live Through This" foi lançado e recebeu ótimas críticas após o suicídio de Cobain, disseram que planejam se reunir e continuar.

"Nós estamos profundamente angustiados com a perda de uma música extremamente talentosa, de grande alma e uma grande amiga" disseram Love, o guitarrista Eric Erlandson e o baterista Patty Schemel em uma declaração dada pela Geffen.

Após o suicídio de Cobain, supostamente depois de ter retornando a usar heroína, Love lamentou a facilidade de se conseguir cocaína em Seattle. Pfaff é a última em uma série de mortes de roqueiros alternativos de Seattle. Andrew Wood, vocalista da Mother Love Bone, morreu de overdose de heroína em 1990 logo após ao lançamento do álbum de debut da banda álbum. (Dois membros do grupo continuaram para formar o Pearl Jam.) Stefanie Sargent, a guitarrista da banda feminina 7 Year Bitch, também morreu de overdose de heroína em 1992, enquanto que Mia Zapata da banda Gits morreu no ano passado em um caso ainda não resolvido.

Pfaff foi criada em Buffalo e estudou piano clássico. Aos 22 ela se apaixonou pelo rock. Quando morava em Minneapolis, ela fundou a banda Janitor Joe e depois entrou para a The Hole in 1993. "Eu já era fã de rock desde adolescente," contou ao jornalista do Times, Robert Hilburn, em março.

Nota do Editor: A coluna Sunday Calendar Pop Eye inclui uma seção sobre Courtney Love usando computador para se comunicar com os fãs. A seção foi impressa antes das notícias sobre a morte da companheira de banda, Kristen Pfaff, ficassem públicas.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

2.

(46908)<DOCNO> LA040994-0049

ROQUEIRO GRUNGE KURT COBAIN MORRE EM APARENTE SUICÍDIO; TRAGÉDIA: UM TIRO TERMINA COM A VIDA DO LÍDER DO NIRVANA. ELE ERA VISTO COMO A VOZ DE UMA GERAÇÃO À DERIVA.

Kurt Cobain, que com sua banda Nirvana deu forma ao rock moderno a partir da alienação impaciente da juventude de hoje, foi encontrado morto sexta-feira, um aparente suicídio por um tiro na cabeça. Ele tinha 27 anos.

O corpo de Cobain foi encontrado por volta das 8:30 da manhã por um eletricista na casa do músico em Seattle. Autoridades disseram que ele estaria morto há um dia; o médico de Seattle que examinou o corpo confirmou a morte de Cobain.

Uma carta de suicídio de uma página foi encontrada perto do corpo. Autoridades não revelaram o conteúdo da nota, mas o eletricista, em entrevista à KNBC-TV, disse que a carta estava coberta de terra de uma planta e escrita em caneta esferográfica – a nota encerrava com as palavras "I love you, I love you." Imediatamente após de saber sobre a morte de Cobain, fãs de sua cidade – os quais ele ajudou a empurrar para a vanguarda do rock 'n' roll dos anos 90 – reuniram-se em frente a sua casa em sinal de luto. Uma chuva fina caía constantemente do céu cinza e frio.

A morte de Cobain chegou enquanto ele lutava contra o uso de drogas e o profundo tumulto pessoal que fazia suas músicas serem tão emocionantes para sua geração. No mês passado, Cobain foi hospitalizado em Roma após entrar em coma por ter consumido álcool de drogas.

Esta semana o Nirvana abandonou uma posição de destaque na próxima "Lollapalooza '94" tour entre rumores de que a banda estava se separando e de problemas com sua saúde.

Com seu cabelo louro e despenteado, estrutura física esbelta, Cobain trouxe seu estilo de rock irascível e explosivo, conhecido como grunge, para o centro da música pop – e sua influência se espalhou para a moda das surradas camisas de flanela, e vestidos que sua vó usaria.

No fim, ele definiu as atitudes da juventude alienada e foi aclamado como a voz de uma geração à deriva.

Talvez a banda tenha ficado mais conhecida por seu álbum "Nevermind" de 1991 e pelo hino sarcástico "Smells like Teen Spirit." O álbum vendeu 10 milhões de cópias em todo o mundo. O álbum seguinte "In Utero," de 1993 também foi um sucesso de crítica e de vendas.

O diretor da Capitol Records, Gary Gersh – que assinou o trio de Cobain com a Geffen Records durante seu tempo no cargo de executivo de artistas e repertório, no final dos anos 80 – se disse estar abalado.

"Este é um dia terrivelmente triste para a música", disse Gersh, que comandou a gravação e o lançamento de "Nevermind" e conhecia Cobain desde 1988.

"Quando as pessoas olharem em retrospecto para a indústria da música, eu acredito que haverá um período pré-Nirvana e um período pós-Nirvana. O álbum 'Nevermind' vai ser encarado como um marco embrionário na história do rock." Notícias das morte de Cobain começaram a ser espalhadas na manhã de sexta-feira, quando o eletricista ligou para uma rádio local e contou que tinha ido à casa do músico para trabalhar em um sistema de segurança. O eletricista disse que olhou pelas grandes janelas do chalé acima da garagem e viu um corpo. Ele chamou a polícia e depois ligou para a rádio.

Mais tarde, uma declaração do médico do condado de King, "Uma autópsia mostrou que Kurt Cobain morreu de um ferimento na cabeça causado por um tiro de espingarda, e que neste momento parece ter sido ele próprio que atirou." Cobain havia admitido anteriormente dificuldades com o uso de heroína, mas insistia que havia se recuperado em 1992 depois de ter tido com sua mulher, a cantora de rock Courtney Love, uma filha, Frances Bean Cobain.

Contudo rumores de recaídas periódicas persistiam. Então no mês passado em Roma, o cantor foi hospitalizado em coma aparentemente causado por uma mistura de sedativos e champagne. Ele se recuperou, mas a bem-sucedida turnê européia do Nirvana estava arruinada. Houve rumores,

na época, que o episódio tratava-se de uma tentativa de suicídio, mas que foram desmentidos por se achar que tratavam-se de detestáveis especulações que sempre parecem cercar estrelas do rock. Secretamente, contudo, aqueles mais próximos ao cantor estavam profundamente preocupados. Seu estado era aparentemente tão grave que fontes contam que seus amigos o confrontaram sobre o problema e tentaram persuadi-lo a procurar tratamento. Alguns dias depois acredita-se que Cobain tenha começado um programa de recuperação para drogados em Los Angeles, mas logo desistiu.

Love, da banda The Hole, estava em Los Angeles quando o corpo foi encontrado sexta-feira e estava sendo confortada por seu advogado. Ela não pode ser contactada para comentar a notícia.

A mãe de Cobain, Wendy O'Conner, falou de sua casa em Aberdeen, Washington, a repórteres que ela não tinha notícias do cantor há seis dias e que estava preocupada. Ele disse que o tinha avisado das vidas problemáticas e mortes prematuras das estrelas do rock. "Eu disse a ele que não se juntasse a este clube de estúpidos." Sua esposa e seus empresários não tinham notícias de seu paradeiro por dias, dizem as fontes.

Críticos comparam a influência musical de Cobain a Bob Dylan e John Lennon. Na sexta-feira muitos aqui traçaram outro paralelo – com o grande guitarrista de Seattle Jimi Hendrix, que morreu aos 27 anos em um apartamento em Londres em 1970 como consequência de uma overdose de drogas.

Cobain trouxe à música não a raiva contra a elite dos anos 60, mas o mal-estar da geração dos anos 90, muitos vindos de lares desfeitos que não tiveram esperança de uma vida melhor e que questionavam o sentido da vida. "Angústia adolescente deu lucro / Agora estou entediado e velho," era a linha de abertura do último álbum do Nirvana.

Kurt Donald Cobain era de uma comunidade rural de Aberdeen, cerca de 100 milhas a sudoeste de Seattle. Ele próprio vindo de uma família desfeita, viveu uma infância infeliz, experiências que quando traduzidas em música tocaram milhões de jovens como ele.

Em uma entrevista segunda-feira com Robert Hilburn do Times, Love tentou descrever a juventude perdida de seu marido. "Existe este lado doce, Jimmy Stewart, 'Mayberry, RFD' nele. Seus programas favoritos de TV são 'Dragnet' and 'Mayberry, RFD' and 'Leave it to Beaver.' Para ele, eles representam a infância perdida," disse ela. E ela falou do futuro: "Eu estou muito contente de ter tido uma menina, mas quero ter um filho com ele também. Assim posso ajudá-lo a compensar pelo relacionamento que ele nunca teve." com o pai dele. Graças ao Nirvana, seguido pelo Pearl Jam e outras bandas de Seattle, esta cidade se transformou a capital do rock contemporâneo, apagando a divisão entre música alternativa e mainstream e trazendo vida nova a um estilo de música que, aos olhos de muitos, tinha tornado-se flácida e comercial.

Quando jovem, Cobain disse a entrevistadores que ele tinha sido influenciado pelos Beatles, depois por Led Zeppelin. Foi a revolução punk dos anos 70 e 80 que foram o catalisador de sua carreira. Com outro músico de Aberdeen, o baixista Krist Novoselic, Cobain formou o Nirvana em 1986. Vivendo de vale alimentação e trabalhando em trabalhos estranhos eles se juntaram a uma série de bateristas. Sucesso foi rápido.

O primeiro álbum do grupo, "Bleach," foi lançado pela Sub Pop, a gravadora revolucionária que foi responsável pelo lançamento do som de Seattle. Com o baterista Dave Grohl, o grupo assinou com Geffen Records' DGC em 1991.

Em 1992, o grupo foi reconhecido como melhor artista novo e melhor vídeo alternativo no MTV Video Music Awards. Quatro vezes foram indicados para Grammys mas nunca ganharam.

As letras de Cobain eram por vezes sensuais e freqüentemente flertavam com temas obscuros. No último verso da música mais famosa do Nirvana "Teen Spirit," Cobain canta: "With the lights out / It's less dangerous. . . ."

Correspondentes Steve Hochman e Chuck Philips em Los Angeles contribuíram para este artigo.

* Histórias Relacionadas: F1 Destino Mortal. A mãe de Kurt Cobain, o problemático vocalista e líder da banda de rock Nirvana, que aparentemente atirou fatalmente contra si próprio sexta-feira, disse ter falado a seu filho para "não se juntar ao clube dos estúpidos" roqueiros que morreram

Appendix B – Relevance Feedback Questionnaire

cedo, muitas vezes por causa de álcool ou drogas. Alguns destes músicos são: * Janis Joplin: Uma cantora carismática e influente, Joplin morreu em 1970 aos 27 anos de overdose de drogas.

* Elvis Presley: Morreu em 1977 em Memphis, Tennessee, aos 42 anos de uma aparente overdose de drogas. A causa da morte consta como parada cardíaca.

* Jimi Hendrix: Um dos guitarristas mais importantes do rock, morreu aos 27 anos em 1970 de overdose

* Donny Hathaway: O cantor e compositor caiu do 15o. andar de um hotel em Manhattan num aparente suicídio em 1979 aos 33anos.

* Jim Morrison: Líder do the Doors, morreu em 1971 de um ataque do coração em Paris. Ele tinha 27 anos.

* Sid Vicious: Baixista da banda de rock punk Sex Pistols, morreu de overdose de heroína em 1979 aos 21 anos.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

1.

(47520)<DOCNO> LA041094-0391

PARA O REGISTRO Kurt Cobain – A história de capa do The Sunday Calendar sobre Courtney Love do grupo de rock The Hole foi impressa antes do aparente suicídio do marido de Love, Kurt Cobain do Nirvana

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

2.

(47519)<DOCNO> LA041094-0390

PARA O REGISTRO Kurt Cobain – A história de capa do The Sunday Calendar sobre Courtney Love do grupo de rock The Hole foi impressa antes do aparente suicídio do marido de Love, Kurt Cobain do Nirvana

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

3.

(68349)<DOCNO> LA070394-0042

OLHO POP: DE 'RUST NEVER SLEEPS' PARA 'SLEEPS WITH ANGELS'

Em sua nota de suicídio, Kurt Cobain citou uma nota famosa de uma canção de Neil Young: "It's better to burn out than to fade away." Young escreveu uma resposta. O veterano cantor e compositor – em vez de um dos colegas grunges de Cobain – irá tornar-se a primeira grande figura do rock a prestar um tributo em uma música ao líder do Nirvana, cuja morte em abril sacudiu o mundo do rock.

"Sleeps With Angels," a música título do álbum de Young que sairá dia 3 de agosto, foi escrita logo após a morte de Cobain em abril. Uma descendente vocalmente escura e desarticulada de "Tonight's the Night" de Young – este clássico de 1974 que também trata de morte e perda – a música fala da confusão e inquietação frequentemente associada com a música e as vidas de Cobain e sua esposa, Courtney Love.

"She was a teen queen She saw the dark side of life He made things happen But when he did it that night He ran up the phone bill She moved around town to town He sleeps with angels He's always on someone's mind He sleeps with angels tonight" * A morte aos 27 anos de Cobain – que assistiu a um dos concertos de Young no Greek Theatre em 1992 com Love – tocou Young profundamente, diz o empresário de Young, Elliot Roberts. "Definitivamente há uma influência em pelo menos uma música e talvez outras (no álbum)," diz Roberts

Há de certo modo, em algumas partes do novo álbum a certeza de que Young está falando a um público mais jovem. A música "Prime of Life" defende o otimismo jovem, enquanto que "Driveby," ao contrário, levanta o espírito de tragédias casuais e inexplicáveis que furam este sentimento de esperança. Young identifica-se com esta geração problemática e inquieta .

O jovial Young, aos 48 anos, tem idade para ser pai de Cobain, e de fato tem sido uma figura paterna para alguns jovens roqueiros. Como um cantor consistentemente independente, vigoroso, ele tem o status de "padrinho do rock alternativo" e tem saboreado chances de trabalhar com aqueles que ele inspirou

Ele escolheu Sonic Youth e Pearl Jam como sessões de abertura em recentes tournês, e foi a possibilidade de encontrar-se com Young que atraiu um relutante Eddie Vedder a ir ao MTV Video Music Awards no último outono. Young tinha aparentemente tentado contactar Cobain por causa de sua overdose quase-fatal em Roma em março, de acordo com Roberts.

"Ele me ligou para ver se nós conseguíamos encontrar Kurt," diz ele. "Eu falei com o empresário de Cobain, tentei descobrir o telefone dele. Mas este acabou sendo o dia da véspera de sua morte e Neil não chegou a entrar em contato com ele. " Steve Hochman

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

4.

(61447)<DOCNO> LA060594-0057

OLHO POP: TALVEZ O ROCK MORRA NESTE MÊS

Alguém – fora os críticos e compositores – realmente ouve as palavras? Joe e John J. Kohut, uma dupla de fãs e compositores da costa leste, aparentemente pensa que a resposta é sim .

De fato, eles acham que os fãs do rock estão tão interessados no que é dito sobre a música que eles organizaram um livro que coleta suas citações favoritas. "Rock Talk" (Faber and Faber: \$13.95) Conta o que os músicos, críticos e outros têm a dizer sobre tudo desde a composição das músicas até a quebra das guitarras.

Aqui estão as 10 Mais do "Rock Talk":

10 -- "No rock você é construído para depois ser desmanchado. Assim como a arquitetura na América, você constrói, deixa estar por uns 10 anos depois chama de desgastado e destrói para propor alguma outra coisa." -- Joni Mitchell

9 -- "O rock pode ser visto como a última tentativa de livrar-se deste universo morto e sem alma e reafirmar o universo da mágica." -- William S. Burroughs

8 -- "No fim, percebi que o rock não era só fama e riqueza. Ao invés, para mim era encontrar meu lugar no mundo, descobrir aonde é o meu lugar." -- Bruce Springsteen

7 -- "Punk é liberdade musical. É dizer, fazer e tocar o que se quer. Nas palavras de Webster, Nirvana significa estar livre da dor, sofrimento e do mundo externo, e isto está bem próximo da minha definição de rock punk." -- Kurt Cobain

6 -- "Longos e lisos espaços de profissionalismo me incomodam. Eu prefiro ter uma banda que pode explodir a qualquer momento. E eu acho que é disto que as pessoas gostam: alma e expressão e fogo e emoção são mais do que perfeitos." -- Neil Young

5 -- "Eu fui tão bem imitado que ouvi pessoas copiarem meus erros." -- Jimi Hendrix

4 -- "Uma das maiores contradições do rock é que ele tem músicas muito pessoais e privadas endereçadas a um público enorme." -- Bono

3 -- "É como uma música de Bob Dylan, que você adorou por anos e anos, e agora ele dá uma entrevista e diz 'Ah, é sobre um cachorro que foi atropelado na rua.' E você pensa 'O que! Aquela música mudou minha opinião sobre a vida por três anos. O que você quer dizer com cachorro morto no meio da rua!'" -- Michael Stipe do R.E.M.

2 -- "Eu abri a porta para várias pessoas, e elas passaram correndo e me deixaram segurando a maçaneta" -- Bo Diddley

1 -- "Vai desaparecer em junho." -- Variety in 1955, sobre o rock 'n' roll -- Robert Levine

Em relação ao tópico, este documento é:

- () relevante
- () irrelevante
- () não sei

5.

(68700) <DOCNO> LA070394-0393

MÚSICA POP: CONTANDO O ESTOQUE DE UM ANO CRUEL PARA O ROCK

O suicídio de Kurt Cobain em abril passado partiu o coração do rock contemporâneo de maneiras que não começaram a ser recuperadas. Uma vida faz sim diferença. Nine Inch Nails e The Hole produziram ótimos álbuns quatro-estrelas durante os últimos seis meses, mas nem estes nem outros lançamentos foram capazes de encher o vazio emocional. Para toda alienação e raiva na música de Cobain, os discos do Nirvana ofereciam um senso de celebração e esperança – mesmo que estas qualidades não fossem fortes o suficiente na vida de Cobain para o salvarem. Com o Nine Inch Nails, Trent Reznor faz música que é amplamente sem esperança – uma análise de desespero que começa onde Cobain parou. Se Cobain ainda estivesse vivo para contrabalançar Nails com sua

música mais relevante, a jornada de Reznor na escuridão teriam sido convincentes o suficiente para atingir uma maior audiência. O álbum, uma obra-prima do gênero intitulado "The Downward Spiral," deveria ter sido um hit.

Com a morte de Cobain, entretanto, o disco tornou-se quase inaudível. Ainda é um trabalho impressionante, mas este é um tempo em que até o mais fã mais escaldado quer um pouco de conforto – e não é surpreendente que o álbum do Nine Inch Nails tenha rapidamente descido nas paradas.

The Hole, o quarteto de Seattle, liderado pela viúva de Cobain, Courtney Love, deparou-se com um problema diferente. Por ela ser conhecida fora do submundo do rock como Sra. Cobain e não por seu próprio talento, o álbum foi encarado com uma certa desconfiança pelo público.

Além disso, a gravadora do The Hole, DGC, foi ineficaz em promover o álbum "Live Through This," porque teria cheirado a exploração. Love também não estava em condições de fazer turnê, então a banda não pode defender-se ao vivo. The Hole foi atingida por outra tragédia no mês passado com a morte da baixista Kristen Pfaff. Após apenas 2 meses, o álbum está fora dos Top 100.

Enquanto os álbuns a seguir permanecem os 10 melhores do ano, pode ser que ainda demore vários meses até que algum álbum (Pearl Jam? Neil Young? R.E.M.?) reacenda a faísca e finalmente ajude a remover a faixa preta que continua a cercar o rock contemporâneo.

* 1 Nine Inch Nails, "The Downward Spiral," (Nothing/TVT/Interscope). Trent Reznor explora o vazio humano e o desespero com destemor assustador nesta coleção que é um marco. "I am the pusher, I am the whore . . . I am the need you have for more," resmunga ele em uma passagem especialmente moderada que focaliza-se nas obsessões e vícios contemporâneos. Em vez de uma celebração ou aceitação do anarquismo ou decadência, o álbum coloca-se como um chamado angustiado por algo em que se acreditar.

* 2 The Hole, "Live Through This," (DGC). Este álbum virtualmente galopa com imaginação e espírito – um trabalho que combina em seus melhores momentos a auto-afirmação de Patti Smith com a intimidade intimacy e candor de Leonard Cohen.

Apesar de bem mais acessível musicalmente do que o disco de estréia da banda em 1992, o Hole não perdeu nem sua individualidade ou audácia. A voz de Love e palavras flutuam entre abrasivas e acariciantes conforme ela explora aspectos do compromisso pessoal e da dúvida pessoal.

* 3 Pavement, "Crooked Rain, Crooked Rain," (Matador). Existem perigos quando uma banda recebe muitas influências (lembra de em Counting Crows), isto pode indicar que as obras do Pavement oferecem ecos de todos de Lou Reed e Gram Parsons até Elliott Murphy e Thurston Moore. Ainda sim, o Pavement não usa as influências para escorar sua música. A trilha aqui afirma um senso convincente de otimismo e descoberta genuinamente jovens.

* 4 Bikini Kill, "Pussy Whipped," (Kill Rock Stars). Anarquia do Noroeste? Esta banda de Olympia, Washington, com seu desacato punk e declarações feministas, relembra o comando revolucionário do Sex Pistols. Da mesma maneira que a clássica banda britânica, Bikini Kill junta dois elementos aparentemente contraditórios. (pura crueza vocal e letras hábeis). O resultado é música que parece inacessível no início, mas acaba pegando como as melhores dos Pistols.

* 5 Beastie Boys, "Ill Communication," (Capitol). Depois de 8 anos de terem provado em "Licensed to Ill" que garotos brancos pode fazer rap, os Beasties – que combinam um pouco da imbecilidade dos Três Patetas com a inteligência de rua do hip-hop – retornam com outro disco extraordinário. O coração do som está em batidas irresistíveis, mas a conversa é divertida também, tratando de tudo desde a música (Lee Dorsey and Bob Marley) à roda (Bobby Knight and Anthony Mason) sem perder um passo.

* 6 Sonic Youth, "Experimental Jet Set, Trash and No Star," (DGC). O problema com alguns veteranos aclamados (ousaríamos a dizer Elvis Costello?) é que o seu trabalho começa a soar forçado, como se fazer música tenha se tornado uma responsabilidade. O milagre destes aclamados veteranos é que eles ainda fazem música que é divertida e alerta.

* 7 Beck, "Mellow Gold," (DGC). "I'm a loser, baby / So why don't you kill me" é um resmungão tão inspirado para os anos 90 quanto "Like a rolling stone" de Dylan foi para os inquietos anos 60

– melhor ainda porque a frase de Beck pode significar qualquer coisa para qualquer pessoa. O resto da união entre folk e hip-hop union não é "Blonde on Blonde," mas Beck tem uma inteligência e coragem excêntrica e questionadora.

* 8 Johnny Cash, "American Recordings," (American). Uma das pedras fundamentais da música pop moderna retorna à vanguarda criativa com um álbum que é tão simples (só voz e violão) quanto é comovente – contos de pecado e salvação comunicados com a intimidade e urgência de alguém que aprendeu bem suas lições.

* 9 The Auteurs, "Now I'm a Cowboy," (Virgin). Luke Haines, o poeta sofredor que guia este grupo inglês, escreve sobre desejos e dúvidas e relacionamentos bizarros com um tipo de sentimentalidade e medo arrebatadores que esteve ausente do coração do rock desde que Bowie o fez nos anos 70.

* 10 Latin Playboys, "Latin Playboys," (Slash/Warner Bros.). Louie Perez e David Hidalgo dos Los Lobos' juntam-se à Mitchell Froom e Tchad Blake para mais sons experimentais de World Beat que fizeram de "Kiko" uma obra tão cheia de sentimentos e vigor. O tom é primariamente divertido e relaxado.

Em relação ao tópico, este documento é:

- ☐ () relevante
- ☐ () irrelevante
- ☐ () não sei

6.

(48417) <DOCNO> LA041494-0393

OC HIGH / NOTÍCIAS E PONTOS DE VISTA DOS ESTUDANTES; O TEMPO VAI DIZER COMO ELE VAI SER LEMBRADO, MAS HOJE NÓS SENTIMOS SAUDADES DELE

Kurt Cobain é agora um membro "do clube". O clube, que referia-se sua mãe, como um grupo seleto de estrelas e ídolos como Jim Morrison, John Lennon e Janis Joplin que tiveram (intencionais ou acidentais) mortes prematuras. Agora, quando o mundo se lembra de Kurt Cobain, eu me pergunto se o seu suicídio vai glorificá-lo ainda mais. Será ele lembrado como um pioneiro, um líder de um movimento alternativo na música de Seattle e a relutante voz de uma juventude problemática, ou como um maníaco depressivo que desistiu da vida e tomou o caminho dos fracos?

Eu sou um grande fã do Nirvana, juntamente com milhões de outros adolescentes pelo mundo. As letras recheadas de Cobain – contos explícitos de raiva, depressão e angústia emocional. – o levaram para as alturas do estrelato imediatamente após o primeiro trabalho do Nirvana "Nevermind," em 1991. Cobain tornou-se a voz relutante da uma geração de adolescentes angustiados dos anos 90. Infelizmente, seu objetivo nunca foi de ter sido famoso. Ele nunca esteve confortável como a voz de uma geração problemática.

"Se existisse um curso de rock para a estrela 101, eu gostaria de fazer, poderia ter me ajudado," disse Cobain em uma entrevista este ano. Seu suicídio foi encarado como uma tragédia, mas não como um choque. O último trabalho do Nirvana, "In Utero," mostrou as letras angustiadas de Cobain, as quais contavam uma história de raiva e depressão. Por exemplo, umas das primeiras frases do álbum explica seu desconforto com a fama : "Teen-age Angst has paid off well. Now I'm bored and old." Houve também rumores que "Rape Me," umas das músicas mais populares do novo álbum, era na verdade um tiro contra um repórter que disse que Cobain e sua esposa (Courtney Love, vocalista do The Hole) estavam usando heroína enquanto Love estava grávida de sua filha, Francis Bean agora com 2 anos de idade.

Anthony DeCurtis, um editor da revista Rolling Stone, disse, "Se Kurt fosse morrer, todos achavam que ia ser de drogas. A coisa que mais perturba é a determinação destrutiva que estava envolvida."

O mundo perdeu um grande artista. Olhando além de seus vários problemas e seu suicídio, Cobain era um homem promissor, jovem, e influente para muitos. Não importa como ele morreu, sua falta irá deixar um grande vazio no cenário da música, e uma geração de adolescentes buscando por uma nova voz.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

7.

(22585)<DOCNO> LA010194-0176

BATIDA POP: TEM CHEIRO DE ROCK CLÁSSICO; O CONCERTO DO NIRVANA NO FORUM DEMONSTRA PORQUE O GRUPO DE SEATTLE PODE SER A BANDA DE ROCK AMERICANA MAIS IMPORTANTE DESDE R.E.M

Kurt Cobain DO Nirvana cantou em alguns momentos na quita a noite no Forum com uma intensidade tão original que ele pareceu totalmente capaz de atos físicos assustadores. Ainda sim houve momentos em que suas letras eram tão vulneráveis que ele parecia absolutamente incapaz de defender-se contra o mais frágil ataque.

É esta habilidade de Cobain de fundir estas emoções conflitantes em suas músicas que fez do álbum do Nirvana "In Utero" o mais obrigatório de 1993 – e estabelece o Nirvana como possivelmente a banda de rock americana mais importante desde R.E.M

Vestindo sua camiseta de pijama e jeans, que já virarm sua marca, Cobain liderou a banda de Seattle na quinta-feira através de uma série de músicas triunfantes – de "Smells Like Teen Spirit" para "Heart-Shaped Box" – que são tão acessíveis e mesmo assim maravilhosamente misteriosas que já tem jeito de clássicos do rock.

Assim como Bob Dylan e Elvis Costello, Cobain prefere letras que são enganadoras – palavras que não só podem ter significados diferentes para pessoas diferentes, como também podem adquirir novos significados a medida em que se muda as próprias atitudes e emoções.

Enquanto ele claramente escreve na posição de um excluído, Cobain as vezes parece dividido sobre a questão básica de tornar-se novamente independente ou de sentir saudade da aceitação daqueles que o excluíram ao longo dos anos.

"Eu sinto falta do conforto de ser triste," diz ele em "Frances Farmer Will Have Her Revenge on Seattle," não dando pistas em sua performance sobre se eles está zombando de seu próprio isolamento ou encontrando consolação nele.

Onde existe muita consciência, literalmente lembra muito o trabalho de Dylan e Costello's work, Cobain escreve de uma maneira mais fragmentada, o que o torna, de um modo estranho, o poeta ideal de esta geração problemática.

Enquanto que aparentemente construídas em torno de idéias e imagens arbitrárias e contrastantes, as canções – com atraentes melodias que relembram os Beatles – estão cheias de alienação e raiva que toca uma audiência extraordinariamente ampla.

Sem perder os fãs punk, que foram os primeiros a reconhecerem a banda, o grupo agora atrai a todos, desde o colegial atualizado até à curiosa maioria, de modo encorajante, um contingente pós-30 descobriu que há algo na música da banda que atinge mais do que só o público da MTV. Em

uma noite rara, quando assentos de festival foram utilizados no Forum, o Nirvana – após sessões de abertura dos veteranos Butthole Surfers e dos Chokebore – exibiram força cativante. O baterista Dave Grohl e o baixista Krist Novoselic tocaram com vigor cru e Pat Smear acompanhou Cobain na guitarra.

O violoncelista Lori Goldston juntou-se ao grupo para versões acústicas de canções como "Polly" e "All Apologies." Nem todos ficaram satisfeitos com as últimas excursões. Durante um dos números acústicos, fãs próximos ao palco interromperam Cobain e jogaram dois sapatos nele. Um sapato bateu em sua cabeça em outro em sua guitarra.

Este é o tipo de público impaciente que o Nirvana deve ignorar, mas houve um outro sinal de frustração que vale prestar atenção. No fim do bis, o grupo fez a velha quebra de guitarra que há tempos faz parte do seu ato.

Se a intenção era um tapa irônico em um ritual do rock ou uma saudação às raízes punk da banda, a tentativa foi tediosa. Em vez de um bônus para a noite, dúzias na multidão entenderam como sinal para ir embora, um benefício para FAIR, uma organização devotada a lutar contra a censura e preconceito na mídia. Enquanto o Nirvana move-se com confiança para um novo ano, esta é uma parte do seu repertório que é melhor ser deixada para trás.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

8.

(59758)<DOCNO> LA052994-0071

OLHO POP: NOTÍCIAS DO NIRVANA:

Foi só uma questão de tempo antes que os rumores comessem sobre o que os membros sobreviventes do Nirvana irão fazer. Aqui vai o primeiro: O baterista Dave Grohl, dizem as fofocas, vai juntar-se ao Pearl Jam, substituindo o supostamente decepcionado Dave Abbruzzese. A resposta em ambos Nirvana e Pearl Jam: Bobagem. "Pearl Jam ainda é uma família feliz," diz uma fonte próxima à banda. "Nenhuma mudança está planejada." (Empresário Kelly Curtis não estava disponível para dar comentários.) Também não houve comentários oficiais do empresário do Nirvana, mas uma fonte próxima à banda disse que nem Grohl nem o baixista Krist Novoselic fizeram planos sobre suas carreiras musicais por causa do suicídio de Kurt Cobain no mês passado. Não se espera que os dois irão formar um novo grupo juntos.

Uma coisa foi planejada por Grohl, que estava de casamento marcado com sua namorada neste fim de semana. Ele vai tocar como parte da banda "Backbeat" no "MTV Movie Awards" sábado. Grohl estava na banda – junto com Mike Mills do R.E.M., Thurston Moore do Sonic Youth, Dave Pirner do Soul Asylum e Greg Dulli do Afghan Whigs' – que gravou músicas para o filme sobre os Beatles. Este evento, a única performance planejada pela banda, será a primeira aparição pública de Grohl desde a morte de Cobain. Steve Hochman

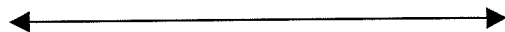
Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

Appendix B – Relevance Feedback Questionnaire

Em uma escala de 1 a 5, como você classifica a dificuldade da tarefa para este tópico (onde 1 significa “muito fácil” e 5 significa “muito difícil”)?

Muito Fácil 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Muito Difícil



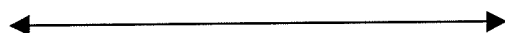
Em uma escala de 1 a 5, o quão confiante você está quanto à precisão dos julgamentos que fez para este tópico (onde 1 significa “muito confiante” e 5 significa “muito inseguro”) ?

Muito Confiante 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Muito Inseguro



Como você considera a sua familiaridade com o tópico da consulta (onde 1 significa “especialista neste assunto” e 5 significa “não tenho nenhum conhecimento do assunto”)?

Especialista 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Nenhum conhecimento



Telefones móveis

Perspectivas sobre o uso de telefones celulares.

Documentos relevantes devem relatar sobre o uso de telefones celulares e o desenvolvimento da indústria de telefones móveis.

1.

(15904)<DOCNO> LA092094-0100

BRIEFLY: TECNOLOGIA & TELECOMUNICAÇÕES

FCC Marca Data para Leilão da Telefonía: Reguladores federais disseram que irão leiloar o primeiro lote de licenças em dezembro para um novo serviço telefónico fornecendo um único número para casa, trabalho ou qualquer outro lugar. Noventa e nove licenças serão disponibilizadas a partir de 5 de dezembro, segundo a Federal Communications Commission. O anúncio dissipou preocupações entre alguns possíveis licitantes de que o leilão poderia terminar sendo adiado até o ano que vem. As licenças, as quais irão permitir companhias a fornecer um serviço de telefonía celular supostamente mais barato e móvel do que os já existentes, irá cobrir 51 áreas geográficas, incluindo Los Angeles-San Diego.

Em relação ao tópico, este documento é:

- ☐ relevante
☐ irrelevante
☐ não sei

2.

(97087)<DOCNO> LA102894-0172

NEXTEL REMOVE OBSTÁCULO PARA A REDE SEM FIO

O plano da Nextel Communications Inc. de criar uma rede nacional de telefones sem fio, fax e pagers removeu um obstáculo regulatório na quinta-feira. A companhia de despachos Rutherford, de New Jersey chegou a um acordo com o departamento de justiça permitindo a aquisição de licenças de rádio da Motorola Inc. Mas Nextel é barrada de comprar novas licenças em mercados onde ela já tem concessões similares de acordo com leis federais.

As licenças são atualmente usadas por ambas as companhias para fornecer serviços de despacho para táxis, caminhões e companhias de entregas. Mas Nextel, com tecnologia digital desenvolvida pela Motorola, planeja utilizar licenças de "rádios móveis especializados" para fornecer serviços semelhantes à telefonía celular, incluindo envio de fax e mensagens.

As companhias planejam competir com as companhias de celular existentes e também com os provedores da próxima geração de telefones celulares e envio de dados conhecidos como serviços de comunicação pessoal.

Em relação ao tópico, este documento é:

- ☐ relevante
☐ irrelevante
☐ não sei

3.

(45798)<DOCNO> LA040494-0089

BREVEMENTE: TECNOLOGIA

Motorola irá Fornecer equipamentos no Japão: A Motorola dos Estados Unidos irá entrar no mercado japonês de telefones pessoais portáteis, uma nova tecnologia de telefones celulares com um maior mercado potencial, Nihon Keizai Shimbun do Tokyo financial daily informa. A Motorola irá fornecer terminais e outros equipamentos para a gigante japonesa das telecomunicações, Nippon Telegraph & Telephone Corp., para serem utilizados em testes de campo que começam neste mês, segundo o jornal. Telefones pessoais portáteis(PHP) são uma tecnologia entre um telefone sem fio e um telefone celular, com não mais alcance do que o sem fio e mais barato do que o celular. Enquanto que menores e mais leves do que os celulares, a principal desvantagem dos PHPs é que eles não podem ser utilizados dentro de um carro ou veículo em movimento. A NEC Corp. e outras companhias japonesas de telecomunicações irão também fornecer seus produtos à NTT no projeto experimental.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

4.

(49000)<DOCNO> LA041694-0274

BREVEMENTE: OUTRAS NOTÍCIAS

BellSouth Corp. pediu a um juiz federal em Washington que sua telefonia celular e outros negócios possam realizar chamadas de longa distância, em um esforço para equiparar-se com competidores como. McCaw Cellular Communications.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

5.

(94372)<DOCNO> LA101794-0202

OPERAÇÃO TERMINA COM FRAUDE DE TELEFONES CELULARES

Confiando em uma operação improvisada, a polícia de Glendale descobriu o que supostamente trata-se de uma das maiores e mais ativas gangues de fraude de telefones celulares do sul da Califórnia, disseram os oficiais no domingo.

Depois de prender o proprietário de uma loja de celulares em Los Angeles na sexta-feira à tarde, a polícia tomou controle das operações da empresa – aprisionando clientes que entravam para comprar telefones ilegais ou para trazer telefones para serem reprogramados, de acordo com fontes familiares com a operação. Além de várias detenções, mais de 50 telefones reprogramados foram apreendidos e milhares de dólares em computadores e equipamentos especializados foram

confiscados, diz Steve Crosby, diretor de assuntos externos da L.A. Cellular, um dos maiores provedores de serviços de telefonia celular de Southland.

Membros da gangue supostamente “clonaram” telefones celulares, um processo pelo qual o telefone é ilegalmente alterado para que as chamadas sejam cobradas nas contas de assinantes legítimos.

O tenente da polícia de Glendale, Ray Edey confirmou a operação mas recusou-se a fornecer maiores comentários antes da entrevista com a imprensa marcada para esta manhã.

Funcionários da L.A. Cellular disseram que "lojas de clonagem" estabeleceram-se como mini companhias telefônicas companhias, cobrando de \$60 a \$100 por mês para serviços ilegais de telefonia celular.

Quando as chamadas ilegais são descobertas, o telefone é desativado. Os clientes então devem retornar à loja para ter o telefone reprogramado para permitir que o serviço ilegal continue.

Oficiais da lei dizem que fraudes em telefones celulares chegaram a proporções epidêmicas no sul da Califórnia, o maior mercado de telefones móveis do país.

Funcionários das companhias de telefonia celular disseram que algumas medidas eletrônicas de segurança sendo implementadas estão somente começando a impedir a clonagem de telefones.

Várias destas medidas ainda não estão em operação.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

6.

(68966)<DOCNO> LA070594-0058

BREVEMENTE: TECNOLOGIA & TELECOMUNICAÇÕES

Companhia alemã busca aliança global na telefonia celular: Deutsche Bundespost Telekom, a maior companhia telefônica da Europa, disse estar com esperanças de tornar-se um dos principais competidores nas comunicações móveis pessoais. A unidade de telefonia móvel da Telekom, DeTeMobil, disse que irá lançar uma rede de telefonia móvel para a região de Moscou ainda este ano, a qual será adicionada a outras quatro licenças digitais e uma analógica na Rússia. Ela disse estar procurando licenças de telefonia móvel na Europa e estudando por volta de 20 opções na Ásia. DeTeMobil quer eventualmente entrar no mercado norte-americano.

A companhia que detém a DeTeMobil's estabeleceu uma aliança com a Sprint Corp. e France Telecom

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

7.

(22366) <DOCNO> LA122994-0309

BREVEMENTE: TECNOLOGIA & TELECOMUNICAÇÕES

Grupo Liderado pela AirTouch Ganha Licença na Espanha: O Grupo Airtel ganhou a licitação para implantar um sistema privado de telefonia celular, que irá competir diretamente com os sistemas existentes operados pela companhia estatal Telefonica, disse Jose Borrel, Ministro de Obras Públicas e Transporte. Airtel inclui a British Telecommunications e parceiros espanhóis como o Banco Santander, Banco Central Hispano, os serviços públicos Union Fenosa e Fecsa e a construtora Cubiertas. AitTouch, sediada em San Francisco, é uma companhia de telefonia celular que foi desligada da Pacific Telesis no início do ano. Airtel tinha dado um lance de \$640 milhões para a licença. Isto faz desta licença a mais cara já leiloadada na Europa.

Em relação ao tópico, este documento é:

- ☐ relevante
- ☐ irrelevante
- ☐ não sei

8.

(18162)<DOCNO> LA102594-0124

GTE IRÁ FORNECER SERVIÇO QUICK-CONNECT 411; TELECOMUNICAÇÕES: CONTRADOS DE COMPANHIAS TELEFÔNICAS COM FIRMAS LOCAIS DE CELULAR IRÃO OFERECER SERVIÇO QUE DISCA O NÚMERO PELO CLIENTE

Com tranquilidade característica, a GTE da Califórnia nos últimos meses conseguiu contratos lucrativos para fornecer serviços de auxílio-à-lista para companhias de telefonia celular no estado.

É um pequeno negócio para a GTE da Califórnia em Thousand Oaks, uma unidade da gigantesca GTE Corp., que é a maior companhia telefônica independente do país, com 20 bilhões de receita anual. Mas serviços via telefonista são geralmente altamente lucrativos para companhias telefônicas locais, e as nossas operações da GTE irão beneficiar-se do crescimento explosivo na telefonia celular. Na Califórnia, a GTE ganhou contratos sobre outros provedores de auxílio-à-lista, incluindo Pacific Bell.

Os novos serviços permitem a um assinante de celular a discar 411 para saber um número de telefone e depois há a opção de ser diretamente conectado ao número pesquisado. Estes serviços conhecidos como "quick-connect" estão sendo introduzidos em todo o país, e eles evitam que os assinantes tenham que terminar uma chamada e discar o número desejado – algo que é particularmente atraente para quem está dirigindo.

Outro atributo do novo serviço 411 para celulares fornecido pela GTE é que quem disca de qualquer lugar dentro da área de Los Angeles não precisa discar o código de acesso, então 555-1212, para obter um número de uma área diferente. Alguém na área de código 818, por exemplo, poderia discar 411 e obter um número de telefone na área de código 714.

GTE está fornecendo estes serviços de auxílio-à-lista para a AirTouch Communications Inc., L.A. Cellular e GTE Mobilnet, a divisão de the celulares da GTE Corp. Isto significa que quando assinantes destas companhias ligam 411, um telefonista da GTE atende e pode fazer a ligação.

GTE ofereceu-se inicialmente a fornecer estes serviços a operadoras de celular na Califórnia e Hawai, mas agora espera expandir o negócio para o país inteiro, disse Jay Burrell, gerente de produto para serviços sem fio da GTE Telephone Operations em Irving, Texas.

O fornecimento de novos serviços para operadoras de celular tem sido fortemente perseguido pela GTE e outras por que a indústria de telefonia celular é prevista a crescer 30% por ano. Além disso, companhias locais cada vez mais são forçadas a procurar novas fontes de renda por causa de

mudanças nas leis e nas tarifas do serviço tradicional, e a provável abertura de mercados locais para a competição das TVs a cabo e companhias telefônicas de longa distância.

Fornecer novos serviços é algo natural para a GTE, disse Burrell. "Nós estamos investindo esforços para sermos os melhores." Outros serviços de 411 que podem ser oferecidos em breve incluem programação de teatro e indicações de como se chegar em algumas destinações, disse Burrell. Também em um futuro próximo, a GTE planeja fornecer todos estes serviços em espanhol.

O negócio de auxílio-à-lista tem boas perspectivas de crescimento. Disse o analista Robert B. Wilkes da Brown Brothers Harriman em Nova Iorque. O anúncio neste mês pela companhia de ligações de longa distância MCI Communications Inc. de seu novo serviço que permite ao assinante discar um número gratuito para poder descobrir qualquer número de telefone publicado nos Estados Unidos ou em outros países. MCI irá oferecer também discar automaticamente para o número sem custo adicional.

Mas 90% das chamadas ao auxílio-à-lista são para chamadas a números locais. E esta atividade deve crescer com o aumento do número de telefones celulares.

A primeira companhia local de telefonia celular a implementar o novo serviço 411 foi GTE Mobilnet, que estreou o serviço em abril na Califórnia, incluindo cinco códigos de área no centro e norte do estado. Todas as chamadas para o 411 são agora automaticamente conectadas ao número pesquisado por 60 centavos a chamada. Uma chamada comum ao auxílio-à-lista sem a conexão automática ao número pesquisado (25 centavos) agora requer a discagem de outro número que não o 411.

David Avirom, diretor de marketing da GTE Mobilnet na Califórnia, disse que, apesar da GTE estar fornecendo os novos serviços para a divisão de celular, licitações de outras companhias também estão sendo consideradas. Avirom disse que utilização da conexão automática entre os assinantes de celular é de quase 100%.

AirTouch começou com seu novo serviço de 411 na área de Los Angeles em junho. A GTE no passado forneceu serviços de 411 para a AirTouch dentro das áreas de chamadas locais servidas pela GTE. Mas isto é "muito diferente do que nós fazíamos antes" disse Melissa May, porta-voz da AirTouch. As mudanças requerem um contrato totalmente novo com a GTE. Apesar da AirTouch ter sido desligada em abril pela empresa proprietária da Pacific Bell, May disse que a AirTouch escolheu a GTE para novos serviços em vez da Pacific Bell porque a GTE ofereceu "melhores serviços a um melhor preço." Uma chamada ao 411 da AirTouch 411 custa 33 centavos, e se o cliente pressiona o botão para ser automaticamente conectado ao número pesquisado, não há custo adicional. Os novos serviços começam em Los Angeles, Orange, Ventura, San Bernardino e Riverside. A companhia espera expandi-los para o país inteiro no início de 1995, quando um assinante de telefone celular poderá discar 411 para um número em qualquer lugar da Califórnia.

A Pacific Bell fornecia anteriormente o serviço de 411 para os celulares de Los Angeles, disse Mike Cymbrowsky, gerente de marketing da L.A. Cellular. Contudo, no início do mês, sob um novo contrato com a GTE, L.A. Cellular também começou a oferecer aos clientes a habilidade de fazer uma ligação de 25 centavos ao 411 e por mais 50 centavos ser conectado ao número pesquisado.

O cliente precisa discar apenas um número na área de serviço da companhia em Los Angeles, Orange, San Bernardino e Riverside.

Burrell da GTE não quis dizer, mas é possível que a GTE também esteja esperando ganhar um contrato para promover serviços semelhantes às demais grandes companhias de celular no estado. A McCaw Cellular Communications Inc., que opera em várias partes da Califórnia como Cellular One, está planejando adicionar um novo serviço de auxílio-à-lista nos próximos meses, disse Sylvia Schnopp, coordenadora de marketing para o distrito da costa sul da Cellular One.

Em relação ao tópico, este documento é:

- ☐ relevante
☐ irrelevante
☐ não sei

9.

(47713)<DOCNO> LA041194-0193

BREVEMENTE: TECNOLOGIA & TELECOMUNICAÇÕES

Secretária Eletrônica de Pager Lançada: Duas companhias revelaram um novo pager em Dallas que atua como uma secretária-eletrônica de bolso. A Motorola Inc. e a Paging Network Inc., as maiores companhias de pagers do país, disseram que esperam começar a anunciar o serviço nacionalmente no ano que vem, cobrando cerca de \$20 por mês. Veja como funciona: A pessoa que envia a mensagem disca um número de telefone especial, então ouve uma saudação gravada pelo receptor e deixa uma mensagem de voz. Esta é enviada por telefone para um terminal de pagers, depois para o satélite e finalmente para o receptor. O receptor ouvirá um bipe ou uma vibração do aparelho.

Em relação ao tópico, este documento é:

- ☐ relevante
☐ irrelevante
☐ não sei

10.

(14888)<DOCNO> LA083194-0219

BREVEMENTE: TECNOLOGIA & TELECOMUNICAÇÕES

Estudo dos Riscos dos Telefones Celulares Revisado: Usuários de telefones celulares podem ser expostos a maiores níveis de energias eletromagnéticas potencialmente nocivas do que acreditava-se anteriormente, de acordo com números revisados por Om Ghandi, um pesquisador da Universidade de Utah. Ele calculou dados de um estudo anterior de 10 telefones celulares e encontrou níveis mais altos de energia eletromagnética do que os relatados anteriormente. Ações na justiça tem alegado uma conexão entre o telefone celular e algumas formas de câncer no cérebro. Um porta-voz da Associação da Indústria de Telecomunicações sediada em Washington disse que, até mesmo utilizando os dados revisados, a quantidade de energia eletromagnética emitida está dentro das regras de segurança estabelecidas.

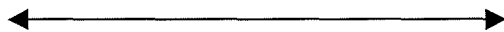
Em relação ao tópico, este documento é:

- ☐ relevante
☐ irrelevante
☐ não sei

Appendix B – Relevance Feedback Questionnaire

Em uma escala de 1 a 5, como você classifica a dificuldade da tarefa para este tópico (onde 1 significa “muito fácil” e 5 significa “muito difícil”)?

Muito Fácil 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Muito Difícil



Em uma escala de 1 a 5, o quão confiante você está quanto à precisão dos julgamentos que fez para este tópico (onde 1 significa “muito confiante” e 5 significa “muito inseguro”) ?

Muito Confiante 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Muito Inseguro



Como você considera a sua familiaridade com o tópico da consulta (onde 1 significa “especialista neste assunto” e 5 significa “não tenho nenhum conhecimento do assunto”)?

Especialista 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Nenhum conhecimento



O que você acharia de um sistema que aceitasse consultas em Português e pesquisasse na Internet documentos em várias línguas (onde 1 é “muito útil” e 5 “sem utilidade para mim”)?

Muito útil 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Sem utilidade



Se estivesse utilizando este sistema, gostaria que os resultados fossem apresentados:

☐ Em sua língua original

☐ Traduzidos para português

Espaço para comentários e sugestões:

APPENDIX C – RELEVANCE FEEDBACK QUESTIONNAIRE (ENGLISH SUBJECTS)

This appendix presents the questionnaire presented to the native English participants of the Relevance Feedback experiment. It is composed by 2 queries and 20 documents.

Searcher 2E

Begin _____

End _____

Instructions

- The task consists in reading the description of a query topic. Imagine that you need to write a report about that topic. You then query a search engine which returns a series of documents. You will then judge each document in relation to the topic in one of three categories "relevant", "irrelevant" or "not sure".
- Documents in which only a part is relevant to the topic should be considered relevant.
- Documents should be judged independently of each other, i.e. the relevance or non relevance of a document should not affect others.
- You will read 2 topics and 10 documents for each topic.
- The documents are news articles published by the Los Angeles Times in 1994.
- There is no time limit.
- The judgements should be made individually.
- You do not need to read each text word for word. Read what you consider enough to make a judgement.

Thanks very much for your participation.

Genetic Engineering

How does genetic engineering effect the human food chain?

Articles must directly address the introduction of genetic engineering, and its effects on the human food chain. They will discuss both pros and cons. Reports on tobacco bioengineering and human gene engineering are not relevant.

1.

(15661)<DOCNO> LA091594-0261 INNOVATION / MICHAEL SCHRAGE: AND ON YOUR LEFT, YOU CAN JUST MAKE OUT THE INSIGHTFUL COLUMNIST GENE The most important cartographers in the world these days don't work at Rand McNally or the CIA. They're not much interested in mountains and rivers and shifting national boundaries. In fact, they don't necessarily know much about traditional geography at all.

No, the influential map makers of the 1990s ply their trade in molecular biology labs, charting the geography of genes. They're rapidly creating a veritable atlas of human, plant and animal genes, figuring out where individual genes are located and what they do.

This is proving to be one of the most productive ventures in the history of cartography: gene mapping is transforming biology, agriculture and medicine as surely as mapping the New World transformed Europe. This year alone, researchers have located more than 45 new disease-causing genes in humans -- just yesterday, scientists announced the discovery of the gene that is believed to cause half of all inherited cases of breast cancer. The better-than-gene-a-week pace has helped spawn the biotech business boom.

But what good is an atlas without a guidebook? Alas, there are no Zagats or Guide Michelins for the genome just yet, no formal tours of the hot spots on uncharted chromosomes. To help rectify this problem, a few experts below point out their favorite sites on the gene map and explain where are the most interesting places to visit: * "The HOX genes," offers Edward P. Lewis, the Thomas Hunt Morgan Professor of Molecular Biology at CalTech. "They are the only way an organism can make a head, thorax and abdomen. They do it by telling the other genes what to do.

That part of the gene has not changed for over 500 million years

* "LINE 1 and LRE1, LRE2 and LRE3," says Maxine Singer, a molecular biologist who is president of the Carnegie Institution of Washington and who has a lab at the National Cancer Institute. "They are places where active copies of the only known copies of human transposable -- so-called jumping genes -- are found.

They're mutagenic when they jump: that is, they cause mutations and disease, like muscular dystrophy, when they move. They appear to be active in early embryonic cells. They're troublemakers, and I'm particularly interested in why these troublemakers have persisted for so long in the genome because of all the trouble they cause

* " Rin ," says William R. Hyatt, the vice president of research and development for Calgene, the company that developed the genetically engineered FlavrSavr tomato. "It's a spot on the tomato locus and mutations map onto there. It's a natural mutant and it inhibits ripening. It's not so much a gene as it is a site. The point is, nobody knows exactly what this gene or locus is -- how it exerts this multitude of effects on the tomato and through what mechanism. We've only recently begun to clone it

* "The sry gene of the Y chromosome," offers Robert L. Sinsheimer, an emeritus professor at UC Santa Barbara and a pioneer in the early days of bacterial virus research. "This was thought to be the central gene for determining maleness and femaleness. . . . There are also corresponding genes on the X (female) chromosome. We thought we understood which genes made you a boy or a girl. It seemed pretty clear and now it's confused

* "The genes responsible for Type 1 diabetes," says venture capitalist Kevin Kinsella, who is also chief executive of La Jolla-based Sequana Therapeutics.

"The sites are chromosomes 6, 11 and 18 and near HLA. There may be as many as five other genes implicated. I'm much more interested now in how gene complexes interact to produce a disease rather than just looking for single genes. The problem is finding all those multiple locations. So I think we're moving into a new era of gene discovery

* "The ACC Synthase genes," suggests David Evans, business development vice president for DNA Plant Technology, which was involved in creating Brown & Williamson's Y-1 high nicotine tobacco. "They oversee the biosynthesis of ethylene. They control ripening and rotting. It's a gene family involved in part of the plant's defense mechanisms against disease. In some fruits, like bananas and tomatoes, there is a direct association between ethylene and fruit rotting. We're looking to suspend that

* "The genes for the NMDA receptor," offers Daniel Koshland, a molecular biologist at UC Berkeley and the editor of Science magazine. "It's the receptor in the brain involved in how we record memory. I'm trying to understand the biochemical shift between short-term memory and long-term memory. We're looking for the gene for it. Where it is on the chromosome is not that important at the moment -- but it may be. I don't have any guarantees that its location is important but that may turn out to give us important insights. Sometimes gene geography can be very important

* "The histone gene clusters," says Laurence Kedes, director of the Institute for Genetic Medicine at USC. "These are the genes that make the proteins that coat the chromosomes. They're the ones responsible for taking the threads of DNA and compacting them into cells. Otherwise, they would be yards long

* MEDICAL BREAKTHROUGH Researchers identify gene responsible for some breast cancers. A1

In relation to the topic, this document is:

- ☐ relevant
☐ irrelevant
☐ not sure

2.

(98683)<DOCNO> LA110394-0361 SEVEN MORE GENETICALLY ALTERED FOODS PASS FDA'S SAFETY TESTS Seven more genetically engineered foods have passed a voluntary safety inspection, the Food and Drug Administration said Wednesday, but scientists questioned whether the FDA is scrutinizing such foods closely enough.

"I am a little troubled," Dr. Marion Nestle, an FDA adviser from New York University, told the agency. "It's as if FDA scientists have accepted these very complicated (safety) assessments on face value

Among the new foods is a potato produced by Monsanto that naturally resists the Colorado potato beetle. The potato produces a natural pesticide that kills beetles that take a bite.

The FDA last spring approved the nation's first biotech food, Calgene Inc.'s Flavr-Savr tomato, which has a gene that allows it to ripen longer on the vine and still get to supermarkets without turning soft.

Now the agency has completed much more cursory safety inspections of the potato and six other genetically altered plants, FDA food safety chief Dr. Jim Maryanski said Wednesday.

Also approved were: * Three more tomatoes, one that works like the Flavr-Savr and two with altered genes that slow the tomatoes' natural ripening, so they can be picked earlier and reach optimal ripeness by the time they hit store shelves.

* A squash genetically altered to resist two deadly viruses.

* Cotton that tolerates a common herbicide, bromoxynil, so farmers could use the herbicide. The cotton is intended to produce cottonseed oil for animal feed.

* Soybean plants that tolerate the herbicide glyphosate. Although the FDA said the plant will be used for animal feed, soybeans and soybean oil are consumed by people.

The FDA is preparing to mandate that it be notified before any genetically engineered food is marketed. That will give the agency a chance to decide what kind of safety review each needs. "We simply don't have the resources to give every plant that intensive review, and we don't see the public health need," Maryanski said.

But scientists warned that these foods must be watched closely to ensure that people with allergies aren't put in danger if, for instance, a nut protein is put into soybeans.

Some environmentalists were also concerned, Reuters reported.

Jeremy Rifkin, head of the Pure Food Campaign, a group opposed to the genetic engineering of food, said: "When you release a genetically engineered plant into the environment, the organisms are alive -- they can reproduce, they can mutate and you cannot recall them to the laboratory

In relation to the topic, this document is:

- ☐ relevant
- ☐ irrelevant
- ☐ not sure

3.

(15689)<DOCNO> LA091594-0289 NEWS BITES: LAST ONE IN THE POOL IS CHICKEN Many of us believe our craving for some foods is so deep it couldn't just be the result of how we were brought up, it's got to be hereditary. Scientists have disagreed, and tended to believe tastes are due to culture or upbringing, but now there is some evidence of a genetic element. Scientists at the University of Cincinnati studied the food preferences of sets of twins and concluded that tastes for certain foods seem to be strongly affected by genetics. The foods that seems to strike a note in the gene pool include orange juice, chicken, cottage cheese and broccoli. BRUCE HENSTELL

In relation to the topic, this document is:

- ☐ relevant
- ☐ irrelevant
- ☐ not sure

4.

(15507)<DOCNO> LA091594-0107 DISCOVERY OF BREAST CANCER GENE CALLED MAJOR ADVANCE; MEDICINE: RESEARCHERS LOCATE DEFECT THAT CAUSES HALF OF INHERITED CASES. EARLY DETECTION MAY SAVE MANY LIVES.

Utah researchers have identified the defective gene believed to cause about half of all inherited cases of breast cancer, a discovery that should make it possible within a year to identify the estimated 600,000 American women who carry the gene.

These women have an 85% risk of developing breast cancer by age 65 and an unusually high risk of developing ovarian cancer. Identifying the women at risk could save many lives by allowing intensive monitoring and early detection of the tumors, which substantially increases the cure rate.

The discovery is the result of a massive effort triggered four years ago when geneticist Mary-Claire King of UC Berkeley showed that the gene, commonly called BRCA1, is located on a small region of chromosome 17, one of the 23 pairs of chromosomes that contain every person's genetic blueprint. The race to locate the gene was ultimately won by a team of more than 45 scientists headed by geneticist Mark H. Skolnick of the University of Utah and Myriad Genetics Inc. of Salt Lake City.

Another team also has identified the approximate location of a second gene, BRCA2, that is believed responsible for most of the remaining cases of inherited breast cancer in women, possibly many sporadic cases of the disease, as well as male breast cancers.

The discovery of BRCA1 is a "major advance (that) will enable us to identify women, particularly younger women, who have an increased risk of developing breast cancer," said molecular biologist Harold Varmus, director of the National Institutes of Health.

Perhaps even more important, however, will be the new understanding the genes provide about how breast cancer occurs, according to Dr. Harmon Eyre, chief medical officer of the American Cancer Society.

"We actually know very little about what causes breast cancer," Eyre said.

"This is a real first step in getting to the bottom of what it is all about

Breast cancer activists were delighted by the discovery, particularly because of its potential implications for understanding the biology of the disease, said Michelle Rakoff, co-chair of the Los Angeles Breast Cancer Alliance. "What we have now is early detection," she said. "What we really want is prevention, and for that we need to know what causes it

Reports on the breakthrough were scheduled to be published next month in the journal Science, but they were released Wednesday after NBC News revealed their contents.

Breast cancer is the most common form of cancer in women, striking an estimated 182,000 U.S. women and 1,000 men each year. It is estimated that one in every nine women will develop the disease sometime in their lives. The five-year-survival rate after treatment for localized breast cancer is 93%, but the rate dips to 72% if the disease has spread in the breast region at the time of diagnosis and to 18% if the disease has spread to other parts of the body.

The disease annually kills about 46,000 women and 300 men.

Researchers have known for more than 100 years that breast cancer runs in families, with this form of the disease accounting for about 5% of all cases.

The inherited form is physically indistinguishable from sporadic breast cancer, but it strikes at an earlier age, with a large proportion of cases occurring in women under the age of 30.

Previous studies of more than 200 families with the BRCA1 gene have revealed that the gene is linked not only to breast and ovarian cancer, but perhaps to colon and prostate cancer as well. Both women and men in these families have an above-normal rate of colon cancer and men have an excessive amount of prostate cancer.

The newly identified gene, Skolnick said, is a very large gene "that had no similarities to any other gene that we know of until the last segment that we found." The structure of that segment suggests that the protein produced from the gene binds to DNA within the cell, but the researchers as yet have no idea what it does after it binds.

Most researchers, however, believe that BRCA1 is a tumor suppressor gene -- one that normally keeps the replication of cells in check. Skolnick's group is working overtime to pinpoint its function when healthy.

"It is going to be difficult, to say the least, to understand how this gene predisposes some women to breast and ovarian cancer," King said. "But no one ever said breast or ovarian cancer would be easy to solve

Skolnick said it would require at least a year to develop a test for the defective gene because researchers need to be sure which mutations cause cancer, which are harmless and which are more likely to cause ovarian cancer rather than breast tumors. "We want to make sure that if you tell a woman she has an elevated risk, it is true," he said.

Based on the size of the gene, a test to detect its presence could cost as much as \$1,000, he said.

Eyre noted that the test will also make it possible to determine when the defective gene is not present. In any family with many women, he said, there is a significant chance that several will develop breast cancer even though they do not carry the gene. In such families, the test should reduce anxiety by showing that unaffected members have no unusual risk for breast cancer, he said.

Currently, some women in families with a history of the disease undergo prophylactic mastectomies to reduce their risk. If the gene is not present, such mastectomies would clearly be unnecessary.

Skolnick's group has also looked for the gene in several patients with sporadic forms of the disease and, to everyone's surprise, was unable to find it. "That's completely unexpected," Eyre said. Generally, he said, a gene that plays a major role in an inherited form of a disease is also important in sporadic cases.

Researchers have higher hopes that the other gene under study, BRCA2, will be linked to sporadic cases once it has been isolated.

The approximate location of BRCA2 on chromosome 13 was determined by a team headed by molecular biologist Douglas F. Easton of the Institute of Cancer Research in London, who studied 15 breast cancer families that do not have BRCA1. "We think that BRCA1 accounts for perhaps half of the inherited cases and BRCA2 accounts for a good fraction of the rest," he said. "But we think there is still another gene to be found

Easton and his colleagues -- a team that includes Skolnick's group -- are now using the same techniques used in Utah to isolate BRCA2 from among the approximately 65 genes located in the small segment of chromosome 13 they identified.

* INNOVATION: Gene mapping is changing biology, farming and medicine. D1

In relation to the topic, this document is:

- ☐ relevant
☐ irrelevant
☐ not sure

5.

(86327)<DOCNO> LA091694-0142 SCIENTISTS UNLOCK 'FANTASTIC' GENE TO BOOST PLANT GROWTH Scientists have unlocked a gene that one day could bring bigger ears of corn, more roses on each bush and even faster-growing forests, all by controlling a hormone inside each plant.

"When I think of the uses -- it's fantastic," said Jędrzej Szerszen, a Michigan State University plant pathologist who helped discover how to genetically manipulate plants' natural growth hormone.

In a study published today in the journal Science, Szerszen and colleagues Robert Bandurski and Krzysztof Szczygłowski isolated and cloned a gene that controls the hormone.

"It's exciting," said Dr. Jerry Cohen of the U.S. Department of Agriculture, who said genetic engineering might begin replacing growth chemicals in two to five years. "This is the first step

Plant growth depends on a hormone called IAA. A gene called iaglu produces an enzyme that controls how much IAA is free within plants to boost growth and how much is inactive, explained Bandurski, who has researched IAA for 30 years.

The Michigan State scientists plucked the iaglu gene out of corn, cloned it and manipulated it to produce more of the vital enzyme.

They then injected tobacco, which grows quickly and accepts genetic engineering well, with the gene. Instead of the normally unwieldy plant, they grew bushier tobacco with more branches -- essentially getting the same amount in a smaller space.

In relation to the topic, this document is:

- ☐ relevant
☐ irrelevant
☐ not sure

6.

(41668)<DOCNO> LA031794-0242 2ND COLON CANCER GENE FOUND; TESTS EXPECTED SOON

With the discovery of a second colon cancer gene, researchers reported Wednesday that they now have identified the causes of more than 90% of the inherited form of the disease.

Between them, the two genes are responsible for one in every six of the 156,000 new cases of colon cancer diagnosed each year. They also appear to account for as many as 30% of sporadic (non-inherited) cases of colon cancer.

Researchers expect within a few months to develop diagnostic tests that will show whether an individual has either gene. If one of the genes is present, doctors can monitor for tumors frequently, enabling their detection while they are still curable by surgery.

"We can reduce cancer deaths in these families by over 90%," said Dr. Bert Vogelstein of Johns Hopkins University, a co-leader of one of the two groups that report the discovery today in the British journal *Nature* and Friday in *Science*.

The discovery could also lead to new anti-cancer drugs within three to five years, predicted microbiologist Richard Fishel of the University of Vermont, a leader of the other team, which reported its findings in *Nature*. "I am very confident that in short order we will be able to develop appropriate therapeutics based on our knowledge of these genes," he said.

The same two teams reported the discovery of the first colon cancer gene in December.

Both genes, which are found on different chromosomes, act like the spell-check function in a word processing program, checking newly synthesized DNA to ensure that no mistakes -- mutations -- occur during cellular proliferation. When either gene is defective, "you accumulate these alterations at an extremely high rate and cancer is the result," said Fishel. Inherited colon cancer usually strikes before the age of 50.

"This is a great triumph for science," said Department of Health and Human Services Secretary Donna Shalala. "These discoveries . . . will lead to screening tests for high-risk individuals soon. Doctors will be able to save countless lives and prevent much needless suffering

Shalala also raised the concern, shared by the researchers, that screening for the genes could sharply impair the ability of carriers -- people who have the genes -- to purchase health insurance. Because virtually everyone who has the genes will develop cancer, carriers might be excluded from coverage. About one in every 200 people has one of the defective genes, Vogelstein said, making it the most common genetic defect in the world.

Both groups said they have tentatively identified a third gene that accounts for most of the remaining 5% to 10% of inherited colon cancer, formally called hereditary nonpolyposis colon cancer. In 1991, other scientists announced the discovery of a gene that causes a rarer form of inherited colon cancer, called familial adenomatous polyposis, that accounts for perhaps 1% of all cases of colon cancer.

Researchers know a great deal about how the two genes work because Fishel and colleague Richard Kolodner, a geneticist at the Dana-Farber Cancer Institute in Boston, have studied virtually identical genes in bacteria and yeast for nearly 20 years. In bacteria, the genes produce two proteins that carry out the "spell-check" function.

The first protein scans newly synthesized DNA to check for any errors that have occurred during its synthesis. If an error is found, the second protein binds to the first and repairs the error.

If the first gene is defective, mistakes in DNA formation are not recognized.

If the second gene is defective, mistakes are not repaired. Either way, mistakes accumulate until a cancer cell is formed.

The researchers are confident an identical process occurs with the human genes, but that has not been proved.

The quick discovery of the human genes, Fishel said in an interview, came about because of the extensive knowledge that had already been compiled about the bacterial genes, a classic example of basic research as opposed to applied.

"This discovery is a coup for the basic research process in the United States," noted Dr. George Vande Woude of the National Cancer Institute.

Both groups are actively developing laboratory tests that could be used to screen members of affected families. That process is complicated, Fishel said, because at least a dozen different mutations of the new gene have been identified and each has to be taken into account in any potential screen.

The test is expected to initially cost about \$1,000, about the same as the colonoscopy now used to search for colon tumors. Such testing will be most appropriate in families where at least three members have been diagnosed with colon or rectal tumors within two generations. Fishel said the testing should be conducted "as early as possible," preferably shortly after birth, because the cancer can strike in the teens. Carriers can also adjust their lifestyle -- adopting a diet low in fat and high in fiber, for example -- to reduce risks.

The cost of the screening is expected to decline with increased use, eventually permitting the test to be applied to the whole population.

There is some evidence that the genes may play a role in other cancers as well, including those of the lungs, intestine, uterus and ovaries. Such tumors occur at a high rate in families with a history of colon cancer, and defective versions of the genes have been identified in some of these tumors from individuals. Gene carriers would also be monitored for the development of these types of tumors.

In relation to the topic, this document is:

- ☐ relevant
- ☐ irrelevant
- ☐ not sure

7.

(82660)<DOCNO> LA090194-0279 SCIENTISTS FIND DEFECTIVE GENE LINKED TO SKIN CANCER Scientists have identified a defective gene that appears to cause an inherited tendency to the deadly skin cancer melanoma and may also play a role in non-inherited melanoma.

In its normal state, the gene acts as a brake on cancer. But people who inherit a defective version apparently lose part of the protection, making them unusually vulnerable to melanoma, the researchers said.

About 32,000 melanoma cases and nearly 7,000 melanoma deaths are expected in the United States this year. Only about 10% of melanoma occurs in people with an inherited tendency and it is unclear what percentage of inherited cases are because of the gene.

Defective versions of the gene may also be involved in many, or maybe even most, cases of non-inherited melanoma, researchers said. In those cases, the gene would be inherited in normal form but later damaged by sunlight or other causes.

Studying the gene may eventually lead to better treatments for the non-inherited disease, researchers said.

More immediately, the discovery of a susceptibility gene could be used to screen for people at risk for melanoma. They could be counseled to take steps such as limiting exposure to the sun, keeping track of possibly precancerous moles and using sunscreen, researchers said.

The normal gene tells the body how to make a protein called p16, which helps regulate cell division. Prior studies suggested the p16 gene discourages tumors. Previous studies also indicated defective versions play a role in cancer.

The new work is reported in the September issue of the journal Nature Genetics by two independent teams of researchers.

One study was done by Nicholas Dracopoli of the National Center for Human Genome Research, part of the National Institutes of Health in Bethesda, Md., and colleagues at NIH and elsewhere.

They found nine of 18 melanoma-prone families they screened showed defects in the p16 gene that appeared related to the disease. In the nine families, the gene was defective in 33 out of 36 people with melanoma.

The second study was done by Alexander Kamb of Myriad Genetics Inc. in Salt Lake City with scientists at the University of Utah and elsewhere.

They studied 13 families who already appeared to have a predisposition to melanoma from some abnormality in the neighborhood of the p16 gene. The researchers found a defective p16 gene in two of the families. They found no evidence of a defective p16 gene in 38 other melanoma-prone families.

Nonetheless, the Utah and NIH studies "really support each other, fundamentally," Kamb said, noting that two types of p16 gene defects appeared in both studies.

In relation to the topic, this document is:

- ☐ relevant
- ☐ irrelevant
- ☐ not sure

8.

(109939)<DOCNO> LA121894-0057 NEW FEAR FROM FOOD SCIENCE: STEALTH ALLERGEN; BIOLOGY: GENETIC ENGINEERING WILL MAKE IT MORE DIFFICULT TO KNOW WHAT REALLY IS INSIDE THE CASSEROLE DISH.

If you're allergic to nuts, it's easy enough to pass up the cocktail mix.

But it may not be so easy for people to pass up some of the potential allergy-triggering mixes from genetic manipulation of corn, tomatoes, squash and countless other foods.

What if, for example, something from a Brazil nut is put into some soybeans, which end up in everything from baby formula to bologna? That kind of question would have seemed foolish before scientists learned to move genetic material from plant to plant, animal to plant, bacteria to plant, virus to plant.

Government regulators, biotech companies and physicians have grown increasingly concerned about the prospect of shuffling allergens into the deck of a food's DNA.

After all, the price of wrongful noshing runs from hives to a life-threatening asthma attack for any of the millions of Americans who are allergic to nuts, eggs, milk peanuts and other foods.

"Anything that you transfer raises a question," said Dr. Hugh A. Sampson, pediatrician and food allergy specialist at the Johns Hopkins University School of Medicine. "You've just got to decide how likely is it to be a problem

Scientists hope that by adding, deleting or changing hereditary traits they can develop crops that taste better, stay fresh longer, offer more nutrients and fewer fats, resist pests, or tolerate herbicides.

But the effort by the world's largest seed company to pack more nourishment into a soybean shows what a risky business genetic engineering can be -- and how well the few safeguards have worked so far.

Researchers at Pioneer Hi-Bred International Inc. of Des Moines knew that Brazil nuts have methionine, a sulfur-rich amino acid that livestock and poultry need.

Soybeans are a top livestock feed, but they are poor in sulfur. Adding methionine would make the soybeans more of a superfeed.

But Pioneer knew about the allergy problem. The Food and Drug Administration told biotech companies in May, 1992, to check for allergens when they test the safety of their products.

Pioneer asked for help from Stephen Taylor, head of the food science department of the University of Nebraska in Lincoln.

Taylor tried the product on blood serum from people with nut allergies and found there was a reaction.

"We are evaluating the research project for the best path forward," said Steve Daugherty, company spokesman.

Options include finding a way to transfer the protein without the allergen, or dropping the project altogether.

The test showed it is possible to screen bio-engineered foods for known allergens. It also holds out the promise that scientists can use biotechnology to start snipping out allergens.

But biotechnology also raises the specter of creating or adding new allergy-inducing substances to the food supply from sources like the viral coat proteins that can help a plant resist disease.

Many scientists argue that the risks are low, particularly since it appears that only a handful of the tens of thousands of proteins out there cause allergies. The new materials introduced by genetic engineering are added in low levels and may not be present in the edible part of the food.

Still, ignorance about food allergies makes it difficult to get as clear a picture of the hazards as most scientists would like.

"There's no foolproof test" for suspected allergens, said Taylor, who nonetheless believes the risks are acceptable.

Scientists estimate that food allergies afflict about 2% of young children and 1% of adults.

Peanuts, milk, eggs, soybeans, fish and shellfish top the list of allergy-causing foods in this country.

The biotechnology revolution could be a boost for food allergy researchers.

"Why would you go out and spend millions of dollars isolating dozens of food allergens when you have no purpose for it?" said Dr. Dean Metcalfe, of the government's National Institute of Allergy and Infectious Diseases. "Suddenly, now, you need it

In relation to the topic, this document is:

- ☐ relevant
- ☐ irrelevant
- ☐ not sure

9.

(112520)<DOCNO> LA123094-0055 SCIENCE / GENETIC RESEARCH; STUDY LINKS ABNORMALITY TO SPINAL MUSCULAR ATROPHY UC Irvine researchers report that they have identified a genetic abnormality they believe causes spinal muscular atrophy, the most common genetic cause of death in infants.

"This is a long-anticipated finding that should quickly improve accuracy in diagnosis of the forms of SMA," said Donald S. Wood, director of science technology for the Muscular Dystrophy Assn. "Most important, it will accelerate our continuing search for a form of treatment for SMA

No treatment now exists.

The disorder, which in its most severe form is uniformly fatal before the age of 2, affects an estimated 20,000 Americans, making it the second most common lethal genetic disorder after cystic fibrosis. It affects one in 10,000 births, but an estimated one in 40 adults carries the defective gene.

The identification of the gene is reported by a team led by UC Irvine molecular biologist John J. Wasmuth in the January issue of the journal Nature Genetics, released today.

Wasmuth's team believes that the disorder is caused by the loss of segments of a gene called XS2G3 on chromosome 5, one of the 23 pairs of chromosomes that make up the human genetic blueprint.

They studied DNA from 29 patients with the most severe form of the disease, called Type I SMA, and compared it to DNA from 235 healthy people. They found that 17 of the 29 patients were missing fragments of XS2G3 in both copies of the gene (one from each parent). In contrast, only two of the 235 healthy people -- both of them known carriers of the disorder -- were missing fragments of the gene, and only in one copy.

The Muscular Dystrophy Assn.'s Wood speculated that fragments are missing in all 29 patients but that researchers were unable to detect them in 12 cases because of the extreme complexity of that area of chromosome 5.

In relation to the topic, this document is:

- ☐ relevant
- ☐ irrelevant
- ☐ not sure

10.

(88274)<DOCNO> LA092394-0150 FOOD PRESERVATIVES FOUND TO BOOST CANCER FIGHTERS Two widely used food preservatives boosted levels of a natural cancer fighter in laboratory animals and appear to do the same thing in humans, a researcher reported Thursday.

Advocates of natural foods have long decried the use of preservatives, but Dr.

Andrew Dannenberg of Cornell Medical College found that the preservatives BHA and BHT "revved up" the gene for an enzyme that helps destroy carcinogens before they trigger tumors.

When the genes are cranked up, they produce more of the enzyme, providing better protection against cancer-causing substances in the environment, Dannenberg reported at the International Conference on Cancer Prevention at Rockefeller University in New York.

BHA and BHT are synthetic additives used as preservatives in cookies, crackers and a wide variety of other foods.

The results do not mean that foods should be pumped up with even more preservatives, Dannenberg said. The findings are important because they uncover a cancer prevention mechanism that appears to be part of the explanation for the well-known anti-cancer properties of broccoli, cauliflower and brussels sprouts.

His research shows that at least part of the effect of those compounds comes from revving up the same gene affected by BHA and BHT.

The gene produces an enzyme with the long name UDP-glucuronosyltransferase, or UGT. The study found elevated levels of the enzyme in the liver, kidneys and small intestines of rats fed higher doses of BHA and BHT than are normally found in foods, Dannenberg said.

He then found preliminary evidence that the substances do the same thing in humans. Dannenberg said he has also found that sulforaphane, an anti-cancer agent recently isolated in broccoli, exerts its action partly in the same way, by revving up the gene for UGT.

The results appear at a time when researchers have become widely convinced of the value of fruits and vegetables in preventing cancer but have made little headway at figuring out exactly why that's true.

One problem is that many people avoid the foods with the most potent anti-cancer effects, said Dr. Walter Willett of the Harvard School of Public Health.


Appendix C – Relevance Feedback Questionnaire (English)

In relation to the topic, this document is:

- () relevant
- () irrelevant
- () not sure


How confident are you about the judgements you made (where 1 means “very confident” and 5 means “very unsure”) ?

Very Confident 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Very unsure



How do you rate your familiarity with this topic (where 1 means “expert on the topic” e 5 significa “no knowledge about the topic”)?

Expert 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ No knowledge



Marriage Jackson-Presley

Find documents that report on the presumed marriage of Michael Jackson with Lisa Marie Presley or on their separation.

In May 1994, the famous pop star, Michael Jackson, was reported to have married Lisa Marie Presley, the daughter of the king of rock and roll. Relevant documents must either contain some details regarding the wedding, such as where or when it was held, or must discuss the later separation of the couple.

1.

(31172) <DOCNO> LA020694-0059

RECORD RACK: IN BRIEF

** Freddie Jackson, "Here It Is," RCA. Could Jackson be the Michael Bolton of black music? Back from a lengthy recording hiatus, Jackson really oversings to his heart's content on several cuts. "Was It Something" and "How Does It Feel" fare well under such treatment, but his overemoting on the love-gone-wrong ballad "My Family" doesn't make sense. Like Bolton, Jackson lacks much in the sincerity department.

New albums are rated on a scale of one star (poor), two stars (fair), three stars (good) and four stars (excellent). CONNIE JOHNSON

In relation to the topic, this document is:

- ☐ relevant
☐ irrelevant
☐ not sure

2.

(49220) <DOCNO> LA041794-0218 MICHAEL JACKSON GETS 'CARING FOR KIDS' AWARD
Michael Jackson may have had his share of legal problems over allegations of child abuse, but 100,000 New York area students have given him a vote of confidence, selecting him to receive a "Caring For Kids" award.

In relation to the topic, this document is:

- ☐ relevant
☐ irrelevant
☐ not sure

3.

(75853) <DOCNO> LA080394-0051

JACKSON-PRESLEY UNION SPARKS SHOCK, DOUBT, LAUGHS; MARRIAGE: CONFIRMATION THAT THE KING OF POP AND THE PRINCESS OF ROCK HAVE INDEED HITCHED CAUSES MANY REACTIONS, WITH SURPRISE TOPPING THE CHART.

Elvis Presley and Michael Jackson fans were united on one thing Tuesday: their shock over the confirmation of marriage of the King of Pop to the daughter of the King of Rock.

"Nobody ever thought they'd see the day when Michael Jackson would come moonwalking down Elvis Presley Boulevard," said Janis Fullilove, a radio talk-show host on WMC in Elvis' hometown of Memphis.

In Jackson's hometown of Gary, Ind., Bill James, who hosts a talk show on WLTH, reported that reaction to the marriage of Jackson and Lisa Marie Presley was skeptical.

"Jackson fans don't believe he's married," James said. "They think it's just for publicity." Bob Job, president of the 367-member New Jersey State Assn. for Elvis, said the marriage poses a quandary. He wonders how he is going to handle the story in Elvis Times, the newsletter he has edited every two weeks for 12 years.

"I'm not going to touch it yet," Job said hesitantly by phone Tuesday. "It's too touchy. You have fans who are fond of Lisa and will go along with anything she does, and others who are not too fond of Michael's publicity, and the trouble he's been getting into with kids.

"Everybody's quiet," Job said of the fan club membership. "They're in shock.

Of course, the other thing on everybody's mind is, where will Graceland fit in way down the road? What will Michael want to do with it?" Cynthia Horner, editorial director for the black teen magazine Right On!, doesn't think Michael married Lisa Marie just to get his gloves on Graceland, or even to improve his negative public image.

"Michael can get publicity anytime he wants," said Horner, who has known the pop star since 1976. "It's not necessary for him to take this route to try and beef up his reputation. He's too private of a person.

"I talked to Michael in April, and at that time he indicated he was looking for a wife, and I told him to look for someone equally as famous and wealthy as he is, so he wouldn't have to worry about someone marrying him just for his money." Upon hearing her advice, Horner reported that Jackson "just smiled." At Tower Records in Hollywood Monday evening, reaction to the Jackson-Presley nuptials was more cynical.

As Calvin Pixley, 37, said: "I don't see love there. Did they get married because they're both famous? I think she really admired (Jackson's) fame -- her father being so famous and all -- and she fell in love with his art, his fame." "It's nonsense," said Darrin Tehrani, 29, a musician who credited both Jackson and Presley among his influences. "Even if it's true, what are their motives? They're two oddballs and this whole thing is a mess. I'm probably going to have bad dreams about this." On KROQ's morning drive-time show, hosts Kevin and Bean spent Tuesday fielding listener theories on how and why the King of Pop and the Princess of Rock got together.

Bean thought Kevin's idea was by far the weirdest.

"He actually thinks that maybe they like each other." Acknowledged Kevin: "It's refreshing to not consider that they're just doing it for the money . . . unlike Tom and Roseanne or Anna Nicole (Smith) and grandpa."

In relation to the topic, this document is:

- ☐ relevant
- ☐ irrelevant
- ☐ not sure

4.

(23290) <DOCNO> LA010594-0010 METROPOLITAN DIGEST / LOS ANGELES COUNTY NEWS IN BRIEF: LOS ANGELES; BOY'S LAWYER SEEKS PHOTOS OF MICHAEL JACKSON'S BODY The attorney representing a 13-year-old boy who claims he was molested by Michael Jackson filed court documents Tuesday in an effort to obtain photographs of the entertainer's body.

Last month, Jackson submitted to a body search by investigators seeking evidence to corroborate the boy's claims.

"We think that the fact that my client can establish what Mr. Jackson looks like naked is very substantial evidence of Mr. Jackson's guilt," said Larry Feldman, the boy's attorney.

Feldman said he filed a motion in court that is a "multiple choice" request: Jackson may provide copies of the police photographs, submit to a second search, or the court may bar the photographs from the civil trial as evidence.

Feldman said he has asked Jackson's attorneys and the Los Angeles County district attorney's office for copies, but they refused.

In relation to the topic, this document is:

- ☐ relevant
- ☐ irrelevant
- ☐ not sure

5.

(35342) <DOCNO> LA022094-0362 MICHAEL APPEARS BRIEFLY AT 'JACKSON FAMILY HONORS' SHOW; ENTERTAINMENT: POP STAR PRESENTS AWARDS TO ELIZABETH TAYLOR AND MOTOWN RECORDS FOUNDER BERRY GORDY.

Billed as a humanitarian event, the "Jackson Family Honors" taping here Saturday had more of the feel of a TV melodrama.

The family's most famous member, beleaguered pop star Michael Jackson, appeared to be a reluctant participant in the more than two-hour production, and the family's most mercurial member, LaToya Jackson, contended that she was barred from the program.

Instead of performing a solo number -- as did sister Janet Jackson, who received a standing ovation early in the evening -- Michael Jackson only agreed to present awards to close friend Elizabeth Taylor and Motown Records founder Berry Gordy.

His only scheduled musical contribution to the show was joining the family in a finale written for the event, which is to be broadcast Tuesday by NBC-TV.

Because the 35-year-old singer's photo had been featured prominently in ads for the reunion show, many in the near-capacity crowd at the 13,000-seat MGM Grand Garden arena had hoped for more music from him.

"The whole reason I paid \$100 for a ticket was to see Michael perform," Julie Marlowe, 24, said. "I thought this was supposed to be a concert." Still, this was the most high-profile U.S. stage appearance by Jackson since he was accused of sexual molestation by a 13-year-old boy in August, and most of the enthusiastic crowd was eager to show its support.

"I've always loved his music, and I wanted to show him we still care about him," said Maria Lopez, 26, after buying a \$100 ticket Saturday afternoon.

Jackson reportedly paid more than \$10 million last month to settle the lawsuit. No criminal charges have been lodged against him.

During the playing of one of his videos to open the 8 p.m. show, the crowd shrieked in approval. The video showed Jackson performing the tune "Leave Me Alone," which was meant to mock his treatment by the tabloid press.

Backstage before presenting his first award, Jackson expressed gratitude for his fans' loyalty.

Appendix C – Relevance Feedback Questionnaire (English)

"Some friends are like shadows, you only see them when the sun shines," he said. "But my fans sustained me even in the dark days. I owe them everything." On a less harmonious note, LaToya Jackson maintained earlier in the week that the family would not let her join the show unless she signed a gag order agreeing not to criticize Michael Jackson or the family any further.

At a morning news conference, Jermaine Jackson, the guiding force behind the TV special, and mother, Katherine, denied that a gag order had been demanded.

Although there had been rumors that LaToya Jackson might try to sneak into the concert in disguise, her husband, Jack Gordon, said Saturday that his wife had rejected the idea after he suggested it.

Despite all the focus on the Jackson family, Saturday's event was not designed as a Jackson reunion, Jermaine Jackson said. Instead, it was planned as an annual event to salute outstanding humanitarians and to raise money for charities.

About \$500,000 is earmarked for charity, mostly for earthquake relief in California, said Jackson family representatives.

In relation to the topic, this document is:

- ☐ relevant
- ☐ irrelevant
- ☐ not sure

6.

(107793) <DOCNO> LA121094-0112 VALLEYWIDE FOCUS: JACKSON FAMILY SUED BY APPAREL MAKER Accusing the Jackson family of "duplicitous" business dealings, attorneys for a French clothing manufacturer filed a lawsuit seeking more than \$17 million, contending that the entertainers reneged on an exclusive deal to market the Jackson family name and logo on a line of jeans and sportswear.

In the suit, the Nelson clothing company of Toulouse, France, alleges that the Jackson family violated a June, 1991, contract granting Nelson "exclusive use of the name and logo of 'the Jackson Family' on any and all products manufactured and distributed" by Nelson.

Named in the suit alleging fraud and breach of contract were Jackson Communications Inc. family patriarch Joseph Jackson, superstars Michael and Janet Jackson, as well as other family members including Tito, Marlon, LaToya and Germaine Jackson.

Attorneys Michael Brush and Michael Sacks contend that after the Jacksons signed the contract giving exclusive rights to Nelson and its president, Tony Elichia, the family formed a Swiss company called United Talents Group and began soliciting other, similar business arrangements.

The attorneys said the Jacksons were paid \$100,000 for licensing rights and additional royalty payments that amounted to hundreds of thousands of dollars.

Many of those payments were in cash, Brush said.

Representatives of the Jackson family could not be reached. However, Jay Lavelly, an attorney for Michael Jackson, said his client "denies he did anything wrongful, fraudulent or improper" and that Michael Jackson never granted the Nelson company the right to use his name or likeness for commercial purposes.

According to a court exhibit attached to the lawsuit, Joseph Jackson wrote a letter in March, 1994, terminating the agreement.

The Jacksons have been hit during the past year with several lawsuits alleging breach of contract in connection with the financially troubled Jackson Family Honors television special. In addition, Michael Jackson paid millions earlier this year to settle a lawsuit filed by a 14-year-old boy who accused Jackson of sexually molesting him. ANN O'NEILL

Appendix C – Relevance Feedback Questionnaire (English)

In relation to the topic, this document is:

- ☐ relevant
☐ irrelevant
☐ not sure

7.

(82283) <DOCNO> LA083094-0190 METRO NEWS: LADERA HEIGHTS; FORMER WIFE OF TITO JACKSON DROWNS IN POOL Delores Jackson, the ex-wife of Michael Jackson's brother Tito, accidentally drowned in a swimming pool while visiting the Ladera Heights home of her boyfriend, the Los Angeles County Sheriff's Department reported.

Jackson, 39, had been swimming with her boyfriend, Donald Bohana, 58, at his home on Holt Avenue about 4 a.m. Saturday. Bohana got out of the pool and left briefly. When he returned, he found Jackson at the bottom of the pool, Sheriff's Deputy Rich Erickson said.

Erickson said Bohana pulled Jackson from the water and called 911. Deputies did not know how long she had been in the water.

Jackson was taken by ambulance to Daniel Freeman Memorial Hospital in Inglewood, where she was pronounced dead.

"Investigators believe Jackson died from accidental drowning, but the exact cause of death will be determined by the medical examiner," Erickson said.

Delores Jackson divorced Tito Jackson in 1988 after 16 years of marriage.

In relation to the topic, this document is:

- ☐ relevant
☐ irrelevant
☐ not sure

8.

(82053) <DOCNO> LA082994-0164 TITO JACKSON'S EX-WIFE DROWNS IN SWIMMING POOL Delores Jackson, the ex-wife of Michael Jackson's brother Tito, accidentally drowned in a swimming pool while visiting the Ladera Heights home of her boyfriend, the Los Angeles County Sheriff's Department reported Sunday.

Jackson, 39, had been swimming with her boyfriend, Donald Bohana, 58, at his home on Holt Avenue about 4 a.m. Saturday. Bohana got out of the pool and left briefly. When he returned, he found Jackson at the bottom of the pool, Sheriff's Deputy Rich Erickson said.

Erickson said Bohana pulled Jackson from the water and called 911. Deputies did not know how long she had been in the water.

Jackson was taken by ambulance to Daniel Freeman Memorial Hospital in Inglewood, where she was pronounced dead.

"Investigators believe Jackson died from accidental drowning, but the exact cause of death will be determined by the medical examiner," Erickson said.

Appendix C – Relevance Feedback Questionnaire (English)

Delores Jackson divorced Tito Jackson in 1988 after 16 years of marriage.

In relation to the topic, this document is:

- ☐ relevant
☐ irrelevant
☐ not sure

9.

(83398) <DOCNO> LA090494-0058 POP EYE AND ONE MORE THING: "I haven't had a penny (from my songs) since 1980," says Neil Innes, in his guise of Ron Nasty, who will front the "New Rutles" in shows Thursday and (newly added) Saturday at the Troubadour. "Now it seems Michael Jackson has bought my publishing by mistake. You haven't got his phone number, have you?" Well, the Rutles were a spoof of the Beatles committed by former Bonzo Dog Band member -- and sometime Monty Python co-conspirator -- Innes with Python's Eric Idle. But this comment isn't so much of a joke. It seems that due to the intentional references to Beatles songs in the tunes Innes wrote for the Rutles, he had to work a deal with ATV, then-owner of the Fab Four song publishing, for some of the royalties. Several years ago, of course, Jackson bought ATV, largely because he wished to own the Beatles songs.

And with it he bought, whether he knew it or not . . . the Rutles canon.

A Rutles spokesman says that Jackson is welcome at the group's shows, and so is his wife, Lisa Marie Presley, for that matter.

In relation to the topic, this document is:

- ☐ relevant
☐ irrelevant
☐ not sure

10.

(75852) <DOCNO> LA080394-0050 WHAT BECOMES A LEGEND MOST: NOTORIETY, MUSIC?; COMMENTARY: THE HOOPLA OVER THE JACKSON-PRESLEY MARRIAGE COULD FURTHER OBSCURE HIS CREATIVE FOCUS. ONLY QUALITY MUSIC-MAKING CAN LEGITIMIZE HIS KING OF POP CLAIM.

And we thought Julia and Lyle were the ultimate odd couple of the '90s.

This week's confirmation of Michael Jackson and Lisa Marie Presley's May wedding takes us beyond the realm of the merely quirky and into the world of the genuinely bizarre.

Here you have Jackson, currently the most famous pop star in the world -- fresh from the humiliation of child molestation charges -- in desperate need of re-establishing his credibility and image.

He marries the daughter of the most famous pop star of all time -- straight from her own heartbreaking divorce -- wondering perhaps if she can ever find someone who understands the pressures of living amid the constant scrutiny that comes with being a Presley.

Now that the wedding has been confirmed, most people can't get enough of the story -- even though it has nothing to do with why we care about Jackson in the first place: his artistry.

The danger for Jackson is that all this public fascination could further obscure his own creative focus and destroy any lingering musical credibility.

In recent years, Jackson has seemed as interested in appearing in the tabloids as on the charts. The funny thing is that the whole story of the marriage was laid out weeks ago in the tabloids but that no one took it seriously.

The reason that the mainstream press laid off the marriage rumor wasn't just that the Jackson advisers all denied the marriage. No one believes celebrities or their advisers anymore anyway.

This was a story that was simply too preposterous.

The King of Pop getting married? To the daughter of the King of Rock? If they made it into a movie, they'd have to put it on the Sci-Fi Channel.

The whole affair gives new meaning to the old Elvis hit "Suspicious Minds." Among the questions raised about this unprecedented pop merger: Is it love? Is it public relations? Is it New Age therapy? Is it a Scientology coup? (Lisa Marie is a follower.) Will Graceland now become an Elvis Presley and Michael Jackson museum? If it is difficult to make sense out of the lives of celebrities who give endless, contradictory interviews (yo, Roseanne!), imagine the challenge of figuring out the motivation of people who don't give any.

How do you begin piecing together the thinking behind this marriage? Once the jokes run out, you begin by realizing they are two of the most private and famous people in the world.

Their experience in the pop spotlight -- regardless of whether Lisa Marie has ever stepped on a stage -- gives them a unique bond.

For Lisa Marie, whose birthright and fortune may sometimes seem as much a burden as a blessing, Michael could stand in her eyes as someone who has spent his life with the same set of riches and emotional strains.

Part of Lisa Marie's unique appeal for Michael is her bloodline. She is one of the few people in the pop world who could make Michael more famous than before -- and you don't dub yourself the King of Pop if you're not obsessed with fame.

Michael believes one way to enhance your fame is to keep people curious. Elvis Presley and Howard Hughes were equally fascinating to him because they were unique . . . mysterious. One reason Michael has given for not talking to the press is, "Elvis didn't do interviews." But there is a danger in such emphasis on fame -- and Michael's work has suffered from it in recent years. He seems to think that fame, rather than the quality of the music, is what ultimately matters.

Whereas there was a sense of revelation and passion in 1982's phenomenally successful "Thriller" album, his two subsequent albums were more commercially calculated than inspired.

The albums both seemed designed to protect his market share by appealing to various radio formats and pop genres. No wonder sales slipped dramatically in the United States. Too much of the music seemed artificial and cold.

It was unreasonable to think that Jackson could ever top the 20 million-plus sales of "Thriller" in the United States, but it was not out of line to expect that he could give us another album with equal heart and fire.

What people remember about Elvis isn't how many No. 1 records he had but his influence on pop music. Elvis didn't invent rock 'n' roll, but he defined it in ways that touched everyone from the Beatles to U2.

Similarly, it's the music that should matter in Jackson's case -- regardless of the inevitable fascination with his public life.

People close to Jackson are ecstatic about the new tracks he has made for the "greatest hits" album due around Nov. 1. They say it is some of the strongest music since "Thriller." But these are some of the same people who denied for weeks the marriage rumors.

In the end, it's the quality of Jackson's music that will determine whether the title of King of Pop is anything more than an empty marketing slogan.

Appendix C – Relevance Feedback Questionnaire (English)

In relation to the topic, this document is:

- () relevant
- () irrelevant
- () not sure

How confident are you about the judgements you made (where 1 means “very confident” and 5 means “very unsure”) ?

Very Confident 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Very unsure



How do you rate your familiarity with this topic (where 1 means “expert on the topic” e 5 significa “no knowledge about the topic”)?

Expert 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ No knowledge



APPENDIX D – CLEF OFFICIAL JUDGEMENTS

This appendix contains the official CLEF relevance judgements for the documents presented to the participants of the Relevance Feedback experiment presented in Chapter 5.

Topic 1

Rank	Docno	Relevance
1	LA100894-0025	Relevant
2	LA042894-0255	Relevant
3	LA120594-0182	Relevant
4	LA101994-0006	Not Relevant
5	LA111194-0058	Relevant
6	LA051994-0012	Relevant
7	LA101394-0365	Relevant
8	LA092594-0179	Not Relevant
9	LA101194-0223	Not Relevant
10	LA100994-0279	Not Relevant

Topic 4

Rank	Docno	Relevance
1	LA020694-0059	Not Relevant
2	LA041794-0218	Not Relevant
3	LA080394-0051	Relevant
4	LA010594-0010	Not Relevant
5	LA022094-0362	Not Relevant
6	LA121094-0112	Not Relevant
7	LA083094-0190	Not Relevant
8	LA082994-0164	Not Relevant
9	LA090494-0058	Not Relevant
10	LA080394-0050	Relevant

Topic 2

Rank	Docno	Relevance
1	LA071994-0037	Relevant
2	LA102594-0045	Relevant
3	LA052494-0107	Relevant
4	LA102994-0154	Not Relevant
5	LA072294-0068	Relevant
6	LA052894-0147	Relevant
7	LA061594-0027	Not Relevant
8	LA052194-0202	Relevant
9	LA062394-0406	Relevant
10	LA052994-0317	Relevant

Topic 5

Rank	Docno	Relevance
1	LA061894-0109	Relevant
2	LA040994-0049	Relevant
3	LA041094-0391	Relevant
4	LA041094-0390	Relevant
5	LA070394-0042	Relevant
6	LA060594-0057	Not Relevant
7	LA070394-0393	Relevant
8	LA041494-0393	Relevant
9	LA010194-0176	Not Relevant
10	LA052994-0071	Not Relevant

Topic 3

Rank	Docno	Relevance
1	LA091594-0261	Relevant
2	LA110394-0361	Relevant
3	LA091594-0289	Not Relevant
4	LA091594-0107	Not Relevant
5	LA091694-0142	Relevant
6	LA031794-0242	Not Relevant
7	LA090194-0279	Not Relevant
8	LA121894-0057	Relevant
9	LA123094-0055	Not Relevant
10	LA092394-0150	Relevant

Topic 6

Rank	Docno	Relevance
1	LA092094-0100	Relevant
2	LA102894-0172	Relevant
3	LA040494-0089	Not Relevant
4	LA041694-0274	Not Relevant
5	LA101794-0202	Not Relevant
6	LA070594-0058	Relevant
7	LA122994-0309	Relevant
8	LA102594-0124	Relevant
9	LA041194-0193	Not Relevant
10	LA083194-0219	Not Relevant

APPENDIX E – DATA FROM USER EXPERIMENTS

This appendix contains the data collected from the user experiment described in Chapter 5 and the statistical tests carried out. The statistical tests were performed on SPSS v. 11.

E1 Variable Description

Here is a list of the variables analysed and their meaning:

- Searcher – identifies the participant
- Qno – identifies the query
- System – the way in which the documents have been presented for that query
- Not Sure – number of times the participant selected the option “not sure” for that query
- Difficulty – level of difficulty of the task for that query (1=very easy and 5=very difficult)
- Confidence – level of confidence in the accuracy of the judgements made (1=very sure and 5=very unsure)
- Knowledge – prior knowledge of the subject of the query (1=expert and 5=no knowledge)
- Missed – percentage of missed relevant, e.g. if the query had 6 relevant documents and the participant only found 3, then missed =0.5
- False – percentage of false alarm
- Sum – sum of missed relevant and false alarm
- Overlap – level of agreement between the participant’s and the official judgements
- Initial Precision – performance of the query before the RF process (using the residual collection)
- Final Precision – performance of the query after the RF process (using the residual collection)
- Change – change in precision achieved through the RF process

E2 Data for all queries

E2.1.1 Main Experiment – Portuguese speakers with poor knowledge of English

Searcher	Qno	System	Not sure	Difficulty	Confidence	Knowledge	Missed	False	Sum	Overlap	Initial Precision	Final Precision	Change
1A	1	Original	10	5	5	5	1.00	0.00	6	0.00	46.33	46.33	0.00
1A	2	Original	10	5	5	5	1.00	0.00	8	0.00	26.65	26.65	0.00
1A	3	MT	0	4	3	5	0.60	0.00	3	0.40	10.64	12.65	0.19
1A	4	MT	0	1	1	2	0.50	0.00	1	0.50	5.99	17.65	1.95
1A	5	HT	0	1	1	2	0.71	0.00	5	0.29	16.61	19.52	0.18
1A	6	HT	0	1	2	3	0.60	0.60	6	0.25	46.94	65.29	0.39
1B	1	Original	4	5	3	4	0.83	0.25	6	0.14	46.33	72.77	0.57
1B	2	Original	2	5	3	2	0.63	0.50	6	0.33	26.65	5.22	-0.80
1B	3	MT	1	4	3	4	0.60	0.40	5	0.29	10.64	1.16	-0.89
1B	4	MT	1	3	4	4	0.50	0.38	4	0.20	5.99	14.37	1.40
1B	5	HT	2	4	3	4	0.71	0.33	6	0.25	16.61	18.70	0.13
1B	6	HT	1	3	3	3	0.60	0.60	6	0.25	46.94	55.18	0.18
1C	1	Original	10	5	1	4	1.00	0.00	6	0.00	46.33	46.33	0.00
1C	2	Original	8	5	1	5	0.88	0.50	8	0.00	26.65	2.73	-0.90
1C	3	MT	0	2	2	4	0.60	0.80	7	0.22	10.64	0.39	-0.96
1C	4	MT	0	3	3	3	0.50	0.63	6	0.14	5.99	5.70	-0.05
1C	5	HT	0	3	2	4	0.57	1.00	7	0.30	16.61	17.48	0.05
1C	6	HT	0	3	3	5	0.20	0.80	5	0.44	46.94	46.52	-0.01
2A	1	Original	2	4	3	4	0.33	0.50	4	0.50	46.33	68.42	0.48
2A	2	Original	10	5	5	5	0.88	0.50	8	0.11	26.65	3.02	-0.89
2A	3	MT	0	2	3	4	0.00	0.20	1	0.83	10.64	28.00	1.63
2A	4	MT	0	2	2	2	0.00	0.00	0	1.00	5.99	18.13	2.03

Appendix E – Data from User Experiments

Searcher	Qno	System	Not sure	Difficulty	Confidence	Knowledge	Missed	False	Sum	Overlap	Initial Precision	Final Precision	Change
2A	5	HT	0	2	2	1	1.00	0.00	7	0.00	16.61	16.61	0.00
2A	6	HT	0	3	3	3	0.40	0.80	6	0.33	46.94	61.13	0.30
2B	1	Original	2	3	3	3	0.17	1.00	5	0.50	46.33	73.25	0.58
2B	2	Original	1	3	3	3	0.63	0.50	6	0.33	26.65	7.84	-0.71
2B	3	MT	0	3	4	3	0.60	0.40	5	0.29	10.64	0.58	-0.95
2B	4	MT	0	1	1	2	0.00	0.00	0	1.00	5.99	18.13	2.03
2B	5	HT	3	3	2	4	0.29	0.00	2	0.71	16.61	27.79	0.67
2B	6	HT	1	3	2	2	0.40	0.80	6	0.33	46.94	44.67	-0.05
2C	1	Original	10	5	5	5	1.00	0.00	6	0.00	46.33	46.33	0.00
2C	2	Original	10	4	5	5	1.00	0.00	8	0.00	26.65	26.65	0.00
2C	3	MT	5	4	4	5	0.60	0.20	4	0.33	10.64	0.77	-0.93
2C	4	MT	3	5	5	5	1.00	0.38	5	0.00	5.99	2.20	-0.63
2C	5	HT	4	4	4	4	0.71	0.67	7	0.22	16.61	19.13	0.15
2C	6	HT	5	4	5	5	0.80	0.60	7	0.13	46.94	52.44	0.12
3A	1	Original	10	5	5	5	1.00	0.00	6	0.00	46.33	46.33	0.00
3A	2	Original	10	5	5	5	1.00	0.00	8	0.00	26.65	19.51	-0.27
3A	3	MT	0	4	3	3	0.40	0.20	3	0.50	10.64	28.00	1.63
3A	4	MT	0	4	3	4	0.50	0.00	1	0.50	5.99	10.03	0.67
3A	5	HT	2	3	3	5	0.86	0.00	6	0.14	16.61	18.09	0.09
3A	6	HT	0	3	3	2	0.80	0.40	6	0.14	46.94	31.15	-0.34
3B	1	Original	10	5	5	5	1.00	0.00	6	0.00	46.33	46.33	0.00
3B	2	Original	10	5	5	5	1.00	0.00	8	0.00	26.65	26.65	0.00
3B	3	MT	0	2	4	2	0.60	0.00	3	0.40	10.64	60.12	4.65
3B	4	MT	1	5	5	3	0.00	0.00	0	1.00	5.99	18.13	2.03
3B	5	HT	0	3	3	3	0.14	0.00	1	0.86	16.61	21.73	0.31
3B	6	HT	0	2	2	4	0.40	0.20	3	0.50	46.94	61.67	0.31
3C	1	Original	10	5	5	5	1.00	0.00	6	0.00	46.33	46.33	0.00
3C	2	Original	10	5	1	5	1.00	0.00	8	0.00	26.65	26.65	0.00

Appendix E – Data from User Experiments

Searcher	Qno	System	Not sure	Difficulty	Confidence	Knowledge	Missed	False	Sum	Overlap	Initial Precision	Final Precision	Change
3C	3	MT	0	5	3	4	0.40	0.00	2	0.60	10.64	4.13	-0.61
3C	4	MT	1	4	3	3	0.50	0.13	2	0.33	5.99	16.69	1.79
3C	5	HT	0	3	3	4	0.57	0.67	6	0.33	16.61	20.18	0.21
3C	6	HT	0	5	4	4	0.80	0.40	6	0.14	46.94	28.28	-0.40
4A	1	HT	0	1	2	1	0.00	1.00	4	0.60	46.33	72.53	0.57
4A	2	HT	5	4	5	4	0.88	0.00	7	0.13	26.65	14.19	-0.47
4A	3	Original	10	5	5	5	1.00	0.00	5	0.00	10.64	10.64	0.00
4A	4	Original	10	5	5	5	1.00	0.00	2	0.00	5.99	5.99	0.00
4A	5	MT	5	1	3	5	0.71	0.33	6	0.25	16.61	15.06	-0.09
4A	6	MT	6	3	3	2	0.80	0.40	6	0.14	46.94	28.28	-0.40
4B	1	HT	1	4	4	5	0.33	0.75	5	0.44	46.33	69.67	0.50
4B	2	HT	1	2	2	4	0.25	1.00	4	0.60	26.65	5.84	-0.78
4B	3	Original	10	5	5	5	1.00	0.00	5	0.00	10.64	10.64	0.00
4B	4	Original	9	5	5	5	1.00	0.00	2	0.00	5.99	5.99	0.00
4B	5	MT	0	5	4	4	0.14	0.67	3	0.67	16.61	24.00	0.44
4B	6	MT	0	3	3	4	0.20	0.80	5	0.44	46.94	44.87	-0.04
4C	1	HT	0	3	4	3	0.17	0.50	3	0.63	46.33	72.10	0.56
4C	2	HT	0	3	4	5	0.75	0.00	6	0.25	26.65	16.34	-0.39
4C	3	Original	10	5	5	4	1.00	0.00	5	0.00	10.64	10.64	0.00
4C	4	Original	10	5	5	4	1.00	0.00	2	0.00	5.99	5.99	0.00
4C	5	MT	0	2	2	3	0.71	0.33	6	0.25	16.61	24.03	0.45
4C	6	MT	0	2	2	2	0.40	0.80	6	0.33	46.94	44.16	-0.06
5A	1	HT	0	3	2	3	0.00	0.50	2	0.75	46.33	72.87	0.57
5A	2	HT	1	3	4	4	0.50	1.00	6	0.40	26.65	7.22	-0.73
5A	3	Original	2	5	5	4	0.40	0.00	2	0.60	10.64	47.73	3.49
5A	4	Original	0	5	5	4	0.00	0.00	0	1.00	5.99	18.13	2.03
5A	5	MT	2	3	3	3	0.57	0.00	4	0.43	16.61	26.31	0.58
5A	6	MT	1	4	3	4	0.60	0.40	5	0.29	46.94	64.57	0.38

Appendix E – Data from User Experiments

Searcher	Qno	System	Not sure	Difficulty	Confidence	Knowledge	Missed	False	Sum	Overlap	Initial Precision	Final Precision	Change
5B	1	HT	0	3	3	4	0.33	0.00	2	0.67	46.33	71.12	0.54
5B	2	HT	0	3	2	3	0.50	0.50	5	0.33	26.65	17.29	-0.35
5B	3	Original	10	5	5	4	1.00	0.00	5	0.00	10.64	10.64	0.00
5B	4	Original	10	4	5	4	1.00	0.00	2	0.00	5.99	5.99	0.00
5B	5	MT	0	2	1	1	0.71	0.00	5	0.29	16.61	26.38	0.59
5B	6	MT	0	3	2	3	0.60	0.60	6	0.25	46.94	32.36	-0.31
5C	1	HT	0	1	5	3	0.17	0.75	4	0.56	46.33	72.58	0.57
5C	2	HT	0	1	4	5	0.50	1.00	6	0.40	26.65	7.22	-0.73
5C	3	Original	0	4	4	5	0.60	0.00	3	0.40	10.64	60.12	4.65
5C	4	Original	0	4	4	4	0.50	0.00	1	0.50	5.99	26.17	3.37
5C	5	MT	0	1	1	5	0.71	0.00	5	0.29	16.61	19.52	0.18
5C	6	MT	0	2	5	2	1.00	0.60	8	0.00	46.94	47.11	0.00
6A	1	HT	0	4	3	2	0.67	0.25	5	0.29	46.33	70.70	0.53
6A	2	HT	3	4	3	3	0.50	0.50	5	0.44	26.65	14.02	-0.47
6A	3	Original	4	5	3	4	1.00	0.40	7	0.00	10.64	0.53	-0.95
6A	4	Original	1	3	3	4	1.00	0.38	5	0.00	5.99	4.59	-0.23
6A	5	MT	1	5	5	5	0.43	0.00	3	0.57	16.61	21.08	0.27
6A	6	MT	4	3	3	3	0.80	0.80	8	0.11	46.94	54.47	0.16
6B	1	HT	2	3	4	4	0.33	1.00	6	0.40	46.33	73.21	0.58
6B	2	HT	7	5	5	5	0.75	0.50	7	0.22	26.65	2.57	-0.90
6B	3	Original	4	5	4	4	0.80	1.00	9	0.10	10.64	0.29	-0.97
6B	4	Original	7	5	5	4	0.50	0.25	3	0.25	5.99	10.55	0.76
6B	5	MT	5	5	4	4	0.43	0.33	4	0.50	16.61	21.39	0.29
6B	6	MT	2	3	3	3	0.00	0.60	3	0.63	46.94	48.96	0.04
6C	1	HT	3	4	4	4	0.67	0.50	6	0.25	46.33	69.76	0.51
6C	2	HT	5	5	5	5	0.50	0.50	5	0.44	26.65	8.75	-0.67
6C	3	Original	10	5	5	5	1.00	0.00	5	0.00	10.64	10.64	0.00
6C	4	Original	10	5	5	5	1.00	0.00	2	0.00	5.99	5.99	0.00

Appendix E – Data from User Experiments

Searcher	Qno	System	Not sure	Difficulty	Confidence	Knowledge	Missed	False	Sum	Overlap	Initial Precision	Final Precision	Change
6C	5	MT	0	4	4	5	0.00	0.67	2	0.78	16.61	19.81	0.19
6C	6	MT	0	5	5	5	0.00	1.00	5	0.50	46.94	51.21	0.09
7A	1	MT	1	3	3	3	0.50	0.25	4	0.43	46.33	72.22	0.56
7A	2	MT	1	2	2	2	0.50	0.00	4	0.50	26.65	11.08	-0.58
7A	3	HT	0	3	3	3	0.20	0.80	5	0.44	10.64	1.29	-0.88
7A	4	HT	0	2	2	2	0.00	0.00	0	1.00	5.99	18.13	2.03
7A	5	Original	1	3	5	3	0.43	0.67	5	0.44	16.61	23.06	0.39
7A	6	Original	0	3	3	4	0.20	0.60	4	0.50	46.94	47.48	0.01
7B	1	MT	0	3	3	4	0.50	0.75	6	0.33	46.33	72.08	0.56
7B	2	MT	1	4	4	5	0.38	0.00	3	0.63	26.65	7.15	-0.73
7B	3	HT	0	3	4	4	0.40	0.60	5	0.38	10.64	0.10	-0.99
7B	4	HT	0	3	3	5	0.00	0.25	2	0.50	5.99	19.51	2.26
7B	5	Original	9	5	5	5	1.00	0.00	7	0.00	16.61	16.61	0.00
7B	6	Original	2	5	5	4	0.40	0.40	4	0.50	46.94	45.74	-0.03
7C	1	MT	2	3	3	4	0.67	0.00	4	0.33	46.33	70.02	0.51
7C	2	MT	1	2	3	3	0.50	0.00	4	0.50	26.65	8.93	-0.66
7C	3	HT	1	2	3	2	0.20	0.20	2	0.67	10.64	22.52	1.12
7C	4	HT	1	2	1	2	0.00	0.25	2	0.50	5.99	17.65	1.95
7C	5	Original	10	5	5	4	1.00	0.00	7	0.00	16.61	16.61	0.00
7C	6	Original	10	5	5	5	1.00	0.00	5	0.00	46.94	46.94	0.00
8A	1	MT	2	5	5	5	1.00	0.00	6	0.00	46.33	46.33	0.00
8A	2	MT	1	4	4	4	0.13	0.50	2	0.78	26.65	9.07	-0.66
8A	3	HT	0	3	3	4	0.20	0.00	1	0.80	10.64	5.09	-0.52
8A	4	HT	0	3	3	3	1.00	0.13	3	0.50	5.99	26.17	3.37
8A	5	Original	10	5	5	5	1.00	0.00	7	0.00	16.61	16.61	0.00
8A	6	Original	10	5	5	5	1.00	0.00	5	0.00	46.94	46.94	0.00
8B	1	MT	1	4	4	4	0.33	0.50	4	0.50	46.33	72.03	0.55
8B	2	MT	0	2	1	5	0.25	0.50	3	0.67	26.65	14.46	-0.46

Appendix E – Data from User Experiments

Searcher	Qno	System	Not sure	Difficulty	Confidence	Knowledge	Missed	False	Sum	Overlap	Initial Precision	Final Precision	Change
8B	3	HT	0	3	3	5	0.20	0.80	5	0.44	10.64	1.11	-0.90
8B	4	HT	0	1	1	3	0.50	0.00	1	0.50	5.99	26.17	3.37
8B	5	Original	0	5	5	5	0.43	0.00	3	0.57	16.61	21.08	0.27
8B	6	Original	1	4	5	5	0.80	0.40	6	0.14	46.94	53.31	0.14
8C	1	MT	0	5	5	4	0.50	0.25	4	0.43	46.33	73.93	0.60
8C	2	MT	0	4	4	5	0.50	0.00	4	0.50	26.65	10.60	-0.60
8C	3	HT	0	2	2	3	0.60	0.20	4	0.33	10.64	6.26	-0.41
8C	4	HT	0	1	1	4	1.00	0.00	2	0.00	5.99	5.99	0.00
8C	5	Original	10	5	5	4	1.00	0.00	7	0.00	16.61	16.61	0.00
8C	6	Original	10	5	5	4	1.00	0.00	5	0.00	46.94	46.94	0.00
9A	1	MT	0	1	1	1	0.17	1.00	5	0.50	46.33	72.83	0.57
9A	2	MT	10	5	5	5	1.00	0.00	8	0.00	26.65	26.65	0.00
9A	3	HT	0	3	2	3	0.20	0.60	4	0.50	10.64	1.18	-0.89
9A	4	HT	0	1	1	1	0.50	0.25	3	0.25	5.99	16.72	1.79
9A	5	Original	2	2	1	1	0.43	0.67	5	0.44	16.61	20.36	0.23
9A	6	Original	0	3	5	5	0.80	0.20	5	0.17	46.94	23.79	-0.49
9B	1	MT	0	2	1	2	0.00	1.00	4	0.60	46.33	72.53	0.57
9B	2	MT	0	4	3	4	0.75	0.50	7	0.22	26.65	9.98	-0.63
9B	3	HT	0	3	2	4	0.20	0.80	5	0.44	10.64	1.01	-0.91
9B	4	HT	0	1	1	3	1.00	0.25	4	0.00	5.99	11.76	0.96
9B	5	Original	0	5	2	3	1.00	0.00	7	0.00	16.61	16.61	0.00
9B	6	Original	0	5	3	4	0.00	1.00	5	0.50	46.94	51.21	0.09
9C	1	MT	4	3	3	3	0.50	0.25	4	0.43	46.33	70.56	0.52
9C	2	MT	1	1	1	5	0.25	0.50	3	0.67	26.65	11.09	-0.58
9C	3	HT	0	1	1	3	0.40	0.00	2	0.60	10.64	2.82	-0.73
9C	4	HT	0	1	1	5	0.00	0.13	1	0.67	5.99	16.23	1.71
9C	5	Original	2	3	1	5	0.00	0.33	1	0.88	16.61	21.25	0.28
9C	6	Original	5	5	4	2	0.60	0.20	4	0.33	46.94	69.28	0.48

Appendix E – Data from User Experiments

E2.1.2 Portuguese speakers with good knowledge of English

Searcher	Qno	System	Not sure	Difficulty	Confidence	Knowledge	Missed	False	Sum	Overlap	Initial Precision	Final Precision	Change
1Bi	1	Original	0	4	3	5	0.50	0.00	3	0.50	46.33	68.52	0.48
1Bi	2	Original	0	3	2	5	0.63	0.00	5	0.38	26.65	9.76	-0.63
1Bi	3	MT	0	3	5	5	0.60	0.00	3	0.40	10.64	3.04	-0.71
1Bi	4	MT	0	3	3	4	0.50	0.00	1	0.50	5.99	10.03	0.67
1Bi	5	HT	0	2	1	5	0.86	0.00	6	0.14	16.61	18.09	0.09
1Bi	6	HT	0	1	1	5	0.60	0.00	3	0.40	46.94	60.07	0.28
3Bi	1	Original	4	2	2	2	0.33	0.25	3	0.57	46.33	74.45	0.61
3Bi	2	Original	4	2	2	3	0.38	0.50	4	0.56	26.65	3.26	-0.88
3Bi	3	MT	1	2	2	2	0.20	1.00	6	0.40	10.64	2.97	-0.72
3Bi	4	MT	4	2	2	2	0.50	0.50	5	0.17	5.99	6.49	0.08
3Bi	5	HT	1	2	2	3	0.00	0.67	2	0.78	16.61	20.27	0.22
3Bi	6	HT	1	2	2	2	0.20	0.80	5	0.44	46.94	48.1	0.02
5Bi	1	MT	0	3	3	3	0.33	0.50	4	0.50	46.33	73.81	0.59
5Bi	2	MT	0	3	2	3	0.38	0.00	3	0.63	26.65	11.78	-0.56
5Bi	3	HT	1	2	3	3	0.40	0.20	3	0.50	10.64	28	1.63
5Bi	4	HT	0	3	2	2	1.00	0.00	2	0.00	5.99	4.59	-0.23
5Bi	5	Original	0	3	2	4	0.71	0.00	5	0.29	16.61	26.38	0.59
5Bi	6	Original	0	1	1	1	0.00	1.00	5	0.50	46.94	51.21	0.09
6Bi	1	HT	0	2	2	2	0.17	0.75	4	0.56	46.33	72.95	0.57
6Bi	2	HT	0	2	1	5	0.13	0.50	2	0.78	26.65	12.36	-0.54
6Bi	3	Original	0	2	2	5	0.40	0.60	5	0.38	10.64	5.28	-0.50
6Bi	4	Original	0	2	2	5	0.50	0.38	4	0.20	5.99	16.23	1.71
6Bi	5	MT	0	3	3	5	0.00	0.00	0	1.00	16.61	21.71	0.31
6Bi	6	MT	0	3	3	5	0.20	1.00	6	0.40	46.94	48.46	0.03
7Bi	1	MT	0	2	2	3	0.33	0.25	3	0.57	46.33	74.05	0.60

Appendix E – Data from User Experiments

Searcher	Qno	System	Not sure	Difficulty	Confidence	Knowledge	Missed	False	Sum	Overlap	Initial Precision	Final Precision	Change
7Bi	2	MT	0	2	2	4	0.25	0.50	3	0.67	26.65	6.88	-0.74
7Bi	3	HT	0	2	2	2	0.40	0.00	2	0.60	10.64	33.86	2.18
7Bi	4	HT	0	1	1	3	0.00	0.00	0	1.00	5.99	18.13	2.03
7Bi	5	Original	0	1	1	3	0.00	0.33	1	0.88	16.61	21.25	0.28
7Bi	6	Original	0	1	1	3	0.00	0.80	4	0.56	46.94	51.21	0.09
9Bi	1	MT	6	4	3	3	0.50	0.25	4	0.43	46.33	73.01	0.58
9Bi	2	MT	8	5	5	5	0.75	0.00	6	0.25	26.65	3.02	-0.89
9Bi	3	HT	0	2	3	4	0.60	0.00	3	0.40	10.64	60.12	4.65
9Bi	4	HT	0	2	3	4	0.00	0.00	0	1.00	5.99	18.13	2.03
9Bi	5	Original	0	2	3	4	0.86	0.00	6	0.14	16.61	19.52	0.18
9Bi	6	Original	0	4	4	4	0.60	0.60	6	0.25	46.94	62.91	0.34

Appendix E – Data from User Experiments

E2.1.3 Native English Speakers

Searcher	Qno	Not Sure	Confidence	Knowledge	Missed	False	Sum	Overlap
1E1	1	1	2	2	0.50	0.25	4	0.40
1E1	2	1	1	3	0.50	0.50	5	0.47
1E2	1	2	4	3	0.50	0.00	3	0.47
1E2	2	0	4	3	0.63	0.00	5	0.46
2E1	3	0	1	5	0.40	0.00	2	0.42
2E1	4	0	1	5	0.00	0.00	0	0.67
2E2	3	1	2	4	0.20	0.00	1	1.00
2E2	4	0	1	3	0.00	0.00	0	0.83
3E1	5	0	2	4	0.86	0.00	6	0.11
3E1	6	2	3	4	0.71	0.00	5	0.08
3E2	5	0	2	3	0.57	0.00	4	0.25
3E2	6	2	3	4	0.43	0.20	4	0.38

E3 Comparison among Hand Translation, Machine Translation and Original Texts

E3.1 ANOVA Missed Relevant

The table below shows that there is no significant difference between the hand and machine translation in terms of missed relevant. There is, however, a significant difference between the original texts and the other two groups.

Multiple Comparisons

Dependent Variable: Missed Relevant

Tukey HSD

(I) System	(J) System	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
HT	MT	-.0034	.05687	.998	-.1379	.1312
	Original	-.3196*	.05687	.000	-.4541	-.1850
MT	HT	.0034	.05687	.998	-.1312	.1379
	Original	-.3162*	.05687	.000	-.4507	-.1816
Original	HT	.3196*	.05687	.000	.1850	.4541
	MT	.3162*	.05687	.000	.1816	.4507

Based on observed means.

*. The mean difference is significant at the .05 level.

E3.2 ANOVA False Alarm

The table below shows that there is no significant difference between the hand and machine translation in terms of false alarm. There is, however, a significant difference between the original texts and the other two groups.

Multiple Comparisons

Dependent Variable: False Alarm

Tukey HSD

(I) System	(J) System	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
HT	MT	.0821	.06021	.363	-.0603	.2245
	Original	.2431*	.06021	.000	.1006	.3855
MT	HT	-.0821	.06021	.363	-.2245	.0603
	Original	.1610*	.06021	.022	.0185	.3034
Original	HT	-.2431*	.06021	.000	-.3855	-.1006
	MT	-.1610*	.06021	.022	-.3034	-.0185

Based on observed means.

*. The mean difference is significant at the .05 level.

E3.3 ANOVA Overlap

The table below shows that there is no significant difference between the hand and machine translation in terms of overlap. There is, however, a significant difference between the original texts and the other two groups.

Multiple Comparisons

Dependent Variable: Overlap

Tukey HSD

(I) System	(J) System	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
HT	MT	-.0242	.04596	.859	-.1329	.0846
	Original	.2353*	.04596	.000	.1265	.3440
MT	HT	.0242	.04596	.859	-.0846	.1329
	Original	.2595*	.04596	.000	.1507	.3682
Original	HT	-.2353*	.04596	.000	-.3440	-.1265
	MT	-.2595*	.04596	.000	-.3682	-.1507

Based on observed means.

*. The mean difference is significant at the .05 level.

E3.4 Testing Different User Groups

E3.4.1 Missed Relevant

Multiple Comparisons

Dependent Variable: Missed Relevant

Tukey HSD

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Bilingual-HT	Bilingual-MT	-.0150	.11909	1.000	-.3697	.3397
	Bilingual-Orig	-.0458	.11909	1.000	-.4005	.3088
	English	-.0783	.11909	.995	-.4330	.2763
	Portuguese-HT	-.0989	.09309	.938	-.3761	.1784
	Portuguese-MT	-.1022	.09309	.928	-.3795	.1750
	Portuguese-Orig	-.4189*	.09309	.000	-.6961	-.1416
Bilingual-MT	Bilingual-HT	.0150	.11909	1.000	-.3397	.3697
	Bilingual-Orig	-.0308	.11909	1.000	-.3855	.3238
	English	-.0633	.11909	.998	-.4180	.2913
	Portuguese-HT	-.0839	.09309	.972	-.3611	.1934
	Portuguese-MT	-.0872	.09309	.966	-.3645	.1900
	Portuguese-Orig	-.4039*	.09309	.000	-.6811	-.1266
Bilingual-Orig	Bilingual-HT	.0458	.11909	1.000	-.3088	.4005
	Bilingual-MT	.0308	.11909	1.000	-.3238	.3855
	English	-.0325	.11909	1.000	-.3872	.3222
	Portuguese-HT	-.0531	.09309	.998	-.3303	.2242
	Portuguese-MT	-.0564	.09309	.997	-.3336	.2209
	Portuguese-Orig	-.3731*	.09309	.002	-.6503	-.0958
English	Bilingual-HT	.0783	.11909	.995	-.2763	.4330
	Bilingual-MT	.0633	.11909	.998	-.2913	.4180
	Bilingual-Orig	.0325	.11909	1.000	-.3222	.3872
	Portuguese-HT	-.0206	.09309	1.000	-.2978	.2567
	Portuguese-MT	-.0239	.09309	1.000	-.3011	.2534
	Portuguese-Orig	-.3406*	.09309	.006	-.6178	-.0633
Portuguese-HT	Bilingual-HT	.0989	.09309	.938	-.1784	.3761
	Bilingual-MT	.0839	.09309	.972	-.1934	.3611
	Bilingual-Orig	.0531	.09309	.998	-.2242	.3303
	English	.0206	.09309	1.000	-.2567	.2978
	Portuguese-MT	-.0033	.05614	1.000	-.1705	.1639
	Portuguese-Orig	-.3200*	.05614	.000	-.4872	-.1528
Portuguese-MT	Bilingual-HT	.1022	.09309	.928	-.1750	.3795
	Bilingual-MT	.0872	.09309	.966	-.1900	.3645
	Bilingual-Orig	.0564	.09309	.997	-.2209	.3336
	English	.0239	.09309	1.000	-.2534	.3011
	Portuguese-HT	.0033	.05614	1.000	-.1639	.1705
	Portuguese-Orig	-.3167*	.05614	.000	-.4839	-.1495
Portuguese-Orig	Bilingual-HT	.4189*	.09309	.000	.1416	.6961
	Bilingual-MT	.4039*	.09309	.000	.1266	.6811
	Bilingual-Orig	.3731*	.09309	.002	.0958	.6503
	English	.3406*	.09309	.006	.0633	.6178
	Portuguese-HT	.3200*	.05614	.000	.1528	.4872
	Portuguese-MT	.3167*	.05614	.000	.1495	.4839

Based on observed means.

*. The mean difference is significant at the .05 level.

E3.4.2 False Alarm

Multiple Comparisons

Dependent Variable: False Alarm
Tukey HSD

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Bilingual-HT	Bilingual-MT	-.0900	.12909	.993	-.4745	.2945
	Bilingual-Orig	-.1283	.12909	.955	-.5128	.2561
	English	.1642	.12909	.864	-.2203	.5486
	Portuguese-HT	-.1896	.10092	.497	-.4902	.1109
	Portuguese-MT	-.0909	.10092	.972	-.3915	.2096
	Portuguese-Orig	.0535	.10092	.998	-.2470	.3541
Bilingual-MT	Bilingual-HT	.0900	.12909	.993	-.2945	.4745
	Bilingual-Orig	-.0383	.12909	1.000	-.4228	.3461
	English	.2542	.12909	.438	-.1303	.6386
	Portuguese-HT	-.0996	.10092	.956	-.4002	.2009
	Portuguese-MT	-.0009	.10092	1.000	-.3015	.2996
	Portuguese-Orig	.1435	.10092	.789	-.1570	.4441
Bilingual-Orig	Bilingual-HT	.1283	.12909	.955	-.2561	.5128
	Bilingual-MT	.0383	.12909	1.000	-.3461	.4228
	English	.2925	.12909	.266	-.0920	.6770
	Portuguese-HT	-.0613	.10092	.997	-.3618	.2392
	Portuguese-MT	.0374	.10092	1.000	-.2631	.3379
	Portuguese-Orig	.1819	.10092	.548	-.1187	.4824
English	Bilingual-HT	-.1642	.12909	.864	-.5486	.2203
	Bilingual-MT	-.2542	.12909	.438	-.6386	.1303
	Bilingual-Orig	-.2925	.12909	.266	-.6770	.0920
	Portuguese-HT	-.3538*	.10092	.010	-.6543	-.0533
	Portuguese-MT	-.2551	.10092	.155	-.5556	.0454
	Portuguese-Orig	-.1106	.10092	.929	-.4112	.1899
Portuguese-HT	Bilingual-HT	.1896	.10092	.497	-.1109	.4902
	Bilingual-MT	.0996	.10092	.956	-.2009	.4002
	Bilingual-Orig	.0613	.10092	.997	-.2392	.3618
	English	.3538*	.10092	.010	.0533	.6543
	Portuguese-MT	.0987	.06085	.668	-.0825	.2799
	Portuguese-Orig	.2431*	.06085	.002	.0619	.4244
Portuguese-MT	Bilingual-HT	.0909	.10092	.972	-.2096	.3915
	Bilingual-MT	.0009	.10092	1.000	-.2996	.3015
	Bilingual-Orig	-.0374	.10092	1.000	-.3379	.2631
	English	.2551	.10092	.155	-.0454	.5556
	Portuguese-HT	-.0987	.06085	.668	-.2799	.0825
	Portuguese-Orig	.1444	.06085	.215	-.0368	.3257
Portuguese-Orig	Bilingual-HT	-.0535	.10092	.998	-.3541	.2470
	Bilingual-MT	-.1435	.10092	.789	-.4441	.1570
	Bilingual-Orig	-.1819	.10092	.548	-.4824	.1187
	English	.1106	.10092	.929	-.1899	.4112
	Portuguese-HT	-.2431*	.06085	.002	-.4244	-.0619
	Portuguese-MT	-.1444	.06085	.215	-.3257	.0368

Based on observed means.

*. The mean difference is significant at the .05 level.

E3.4.3 Overlap

Multiple Comparisons

Dependent Variable: Overlap

Tukey HSD

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Bilingual-HT	Bilingual-MT	.0567	.09882	.997	-.2376	.3510
	Bilingual-Orig	.1158	.09882	.904	-.1785	.4101
	English	.0883	.09882	.973	-.2060	.3826
	Portuguese-HT	.1439	.07725	.508	-.0862	.3740
	Portuguese-MT	.1191	.07725	.719	-.1110	.3491
	Portuguese-Orig	.3791*	.07725	.000	.1490	.6091
Bilingual-MT	Bilingual-HT	-.0567	.09882	.997	-.3510	.2376
	Bilingual-Orig	.0592	.09882	.997	-.2351	.3535
	English	.0317	.09882	1.000	-.2626	.3260
	Portuguese-HT	.0872	.07725	.918	-.1428	.3173
	Portuguese-MT	.0624	.07725	.984	-.1677	.2925
	Portuguese-Orig	.3224*	.07725	.001	.0923	.5525
Bilingual-Orig	Bilingual-HT	-.1158	.09882	.904	-.4101	.1785
	Bilingual-MT	-.0592	.09882	.997	-.3535	.2351
	English	-.0275	.09882	1.000	-.3218	.2668
	Portuguese-HT	.0281	.07725	1.000	-.2020	.2581
	Portuguese-MT	.0032	.07725	1.000	-.2268	.2333
	Portuguese-Orig	.2632*	.07725	.014	.0332	.4933
English	Bilingual-HT	-.0883	.09882	.973	-.3826	.2060
	Bilingual-MT	-.0317	.09882	1.000	-.3260	.2626
	Bilingual-Orig	.0275	.09882	1.000	-.2668	.3218
	Portuguese-HT	.0556	.07725	.991	-.1745	.2856
	Portuguese-MT	.0307	.07725	1.000	-.1993	.2608
	Portuguese-Orig	.2907*	.07725	.004	.0607	.5208
Portuguese-HT	Bilingual-HT	-.1439	.07725	.508	-.3740	.0862
	Bilingual-MT	-.0872	.07725	.918	-.3173	.1428
	Bilingual-Orig	-.0281	.07725	1.000	-.2581	.2020
	English	-.0556	.07725	.991	-.2856	.1745
	Portuguese-MT	-.0248	.04658	.998	-.1636	.1139
	Portuguese-Orig	.2352*	.04658	.000	.0964	.3739
Portuguese-MT	Bilingual-HT	-.1191	.07725	.719	-.3491	.1110
	Bilingual-MT	-.0624	.07725	.984	-.2925	.1677
	Bilingual-Orig	-.0032	.07725	1.000	-.2333	.2268
	English	-.0307	.07725	1.000	-.2608	.1993
	Portuguese-HT	.0248	.04658	.998	-.1139	.1636
	Portuguese-Orig	.2600*	.04658	.000	.1213	.3987
Portuguese-Orig	Bilingual-HT	-.3791*	.07725	.000	-.6091	-.1490
	Bilingual-MT	-.3224*	.07725	.001	-.5525	-.0923
	Bilingual-Orig	-.2632*	.07725	.014	-.4933	-.0332
	English	-.2907*	.07725	.004	-.5208	-.0607
	Portuguese-HT	-.2352*	.04658	.000	-.3739	-.0964
	Portuguese-MT	-.2600*	.04658	.000	-.3987	-.1213

Based on observed means.

*. The mean difference is significant at the .05 level.

E4 Comparing Confidence, Difficulty and Prior Knowledge

Multiple Comparisons

Dependent Variable: Difficulty

Tukey HSD

(I) Topic	(J) Topic	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	-.11	.366	1.000	-1.17	.95
	3	.00	.366	1.000	-1.06	1.06
	4	.48	.366	.777	-.58	1.54
	5	.19	.366	.996	-.87	1.24
	6	.07	.366	1.000	-.98	1.13
2	1	.11	.366	1.000	-.95	1.17
	3	.11	.366	1.000	-.95	1.17
	4	.59	.366	.588	-.46	1.65
	5	.30	.366	.966	-.76	1.35
	6	.19	.366	.996	-.87	1.24
3	1	.00	.366	1.000	-1.06	1.06
	2	-.11	.366	1.000	-1.17	.95
	4	.48	.366	.777	-.58	1.54
	5	.19	.366	.996	-.87	1.24
	6	.07	.366	1.000	-.98	1.13
4	1	-.48	.366	.777	-1.54	.58
	2	-.59	.366	.588	-1.65	.46
	3	-.48	.366	.777	-1.54	.58
	5	-.30	.366	.966	-1.35	.76
	6	-.41	.366	.876	-1.46	.65
5	1	-.19	.366	.996	-1.24	.87
	2	-.30	.366	.966	-1.35	.76
	3	-.19	.366	.996	-1.24	.87
	4	.30	.366	.966	-.76	1.35
	6	-.11	.366	1.000	-1.17	.95
6	1	-.07	.366	1.000	-1.13	.98
	2	-.19	.366	.996	-1.24	.87
	3	-.07	.366	1.000	-1.13	.98
	4	.41	.366	.876	-.65	1.46
	5	.11	.366	1.000	-.95	1.17

Based on observed means.

Difficulty

Tukey HSD^{a,b}

Topic	N	Subset
		1
4	27	3.11
5	27	3.41
6	27	3.52
1	27	3.59
3	27	3.59
2	27	3.70
Sig.		.588

Means for groups in homogeneous subsets are displayed.

Based on Type III Sum of Squares

The error term is Mean Square(Error) = 1.811.

a. Uses Harmonic Mean Sample Size = 27.000.

b. Alpha = .05.

Multiple Comparisons

Dependent Variable: Confidence

Tukey HSD

(I) Topic	(J) Topic	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	.00	.374	1.000	-1.08	1.08
	3	.04	.374	1.000	-1.04	1.12
	4	.41	.374	.886	-.67	1.49
	5	.37	.374	.921	-.71	1.45
	6	-.07	.374	1.000	-1.15	1.01
2	1	.00	.374	1.000	-1.08	1.08
	3	.04	.374	1.000	-1.04	1.12
	4	.41	.374	.886	-.67	1.49
	5	.37	.374	.921	-.71	1.45
	6	-.07	.374	1.000	-1.15	1.01
3	1	-.04	.374	1.000	-1.12	1.04
	2	-.04	.374	1.000	-1.12	1.04
	4	.37	.374	.921	-.71	1.45
	5	.33	.374	.948	-.75	1.41
	6	-.11	.374	1.000	-1.19	.97
4	1	-.41	.374	.886	-1.49	.67
	2	-.41	.374	.886	-1.49	.67
	3	-.37	.374	.921	-1.45	.71
	5	-.04	.374	1.000	-1.12	1.04
	6	-.48	.374	.792	-1.56	.60
5	1	-.37	.374	.921	-1.45	.71
	2	-.37	.374	.921	-1.45	.71
	3	-.33	.374	.948	-1.41	.75
	4	.04	.374	1.000	-1.04	1.12
	6	-.44	.374	.843	-1.52	.64
6	1	.07	.374	1.000	-1.01	1.15
	2	.07	.374	1.000	-1.01	1.15
	3	.11	.374	1.000	-.97	1.19
	4	.48	.374	.792	-.60	1.56
	5	.44	.374	.843	-.64	1.52

Based on observed means.

Confidence

Tukey HSD^{a,b}

Topic	N	Subset
		1
4	27	3.07
5	27	3.11
3	27	3.44
1	27	3.48
2	27	3.48
6	27	3.56
Sig.		.792

Means for groups in homogeneous subsets are displayed.

Based on Type III Sum of Squares

The error term is Mean Square(Error) = 1.893.

a. Uses Harmonic Mean Sample Size = 27.000.

b. Alpha = .05.

Multiple Comparisons

Dependent Variable: Knowledge

Tukey HSD

(I) Topic	(J) Topic	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	-.63	.303	.303	-1.50	.24
	3	-.22	.303	.977	-1.10	.65
	4	.15	.303	.996	-.72	1.02
	5	-.07	.303	1.000	-.95	.80
	6	.07	.303	1.000	-.80	.95
2	1	.63	.303	.303	-.24	1.50
	3	.41	.303	.758	-.47	1.28
	4	.78	.303	.111	-.10	1.65
	5	.56	.303	.446	-.32	1.43
	6	.70	.303	.190	-.17	1.58
3	1	.22	.303	.977	-.65	1.10
	2	-.41	.303	.758	-1.28	.47
	4	.37	.303	.824	-.50	1.24
	5	.15	.303	.996	-.72	1.02
	6	.30	.303	.924	-.58	1.17
4	1	-.15	.303	.996	-1.02	.72
	2	-.78	.303	.111	-1.65	.10
	3	-.37	.303	.824	-1.24	.50
	5	-.22	.303	.977	-1.10	.65
	6	-.07	.303	1.000	-.95	.80
5	1	.07	.303	1.000	-.80	.95
	2	-.56	.303	.446	-1.43	.32
	3	-.15	.303	.996	-1.02	.72
	4	.22	.303	.977	-.65	1.10
	6	.15	.303	.996	-.72	1.02
6	1	-.07	.303	1.000	-.95	.80
	2	-.70	.303	.190	-1.58	.17
	3	-.30	.303	.924	-1.17	.58
	4	.07	.303	1.000	-.80	.95
	5	-.15	.303	.996	-1.02	.72

Based on observed means.

Knowledge

Tukey HSD^{a,b}

Topic	N	Subset
		1
4	27	3.52
6	27	3.59
1	27	3.67
5	27	3.74
3	27	3.89
2	27	4.30
Sig.		.111

Means for groups in homogeneous subsets are displayed.

Based on Type III Sum of Squares

The error term is Mean Square(Error) = 1.236.

a. Uses Harmonic Mean Sample Size = 27.000.

b. Alpha = .05.

E5 Relationship Between Judgement Errors and Change in Performance

E5.1 All Data

Correlations

		Not Sure	Difficulty	Confidence	Knowledge	Missed Relevant	False Alarm	M + F	Residual Initial	Final Precision	Change	Overlap
Not Sure	Pearson Correlation	1	.612**	.527**	.438**	.664**	-.387**	.360**	.026	-.104	-.216**	-.646**
	Sig. (2-tailed)		.000	.000	.000	.000	.000	.000	.743	.186	.006	.000
	N	162	162	162	162	162	162	162	162	162	162	162
Difficulty	Pearson Correlation	.612**	1	.705**	.525**	.413**	-.246**	.256**	.074	-.039	-.182*	-.386**
	Sig. (2-tailed)	.000		.000	.000	.000	.002	.001	.350	.623	.020	.000
	N	162	162	162	162	162	162	162	162	162	162	162
Confidence	Pearson Correlation	.527**	.705**	1	.511**	.375**	-.189*	.211**	.101	.030	-.097	-.343**
	Sig. (2-tailed)	.000	.000		.000	.000	.016	.007	.201	.703	.220	.000
	N	162	162	162	162	162	162	162	162	162	162	162
Knowledge	Pearson Correlation	.438**	.525**	.511**	1	.289**	-.178*	.163*	-.018	-.172*	-.246**	-.273**
	Sig. (2-tailed)	.000	.000	.000		.000	.024	.038	.821	.028	.002	.000
	N	162	162	162	162	162	162	162	162	162	162	162
Missed Relevant	Pearson Correlation	.664**	.413**	.375**	.289**	1	-.425**	.583**	-.011	-.170*	-.242**	-.917**
	Sig. (2-tailed)	.000	.000	.000	.000		.000	.000	.894	.030	.002	.000
	N	162	162	162	162	162	162	162	162	162	162	162
False Alarm	Pearson Correlation	-.387**	-.246**	-.189*	-.178*	-.425**	1	.290**	.347**	.209**	-.270**	.177*
	Sig. (2-tailed)	.000	.002	.016	.024	.000		.000	.000	.007	.001	.024
	N	162	162	162	162	162	162	162	162	162	162	162
M + F	Pearson Correlation	.360**	.256**	.211**	.163*	.583**	.290**	1	.379**	.057	-.511**	-.711**
	Sig. (2-tailed)	.000	.001	.007	.038	.000	.000		.000	.470	.000	.000
	N	162	162	162	162	162	162	162	162	162	162	162
Residual Initial	Pearson Correlation	.026	.074	.101	-.018	-.011	.347**	.379**	1	.787**	-.219**	-.099
	Sig. (2-tailed)	.743	.350	.201	.821	.894	.000	.000		.000	.005	.208
	N	162	162	162	162	162	162	162	162	162	162	162
Final Precision	Pearson Correlation	-.104	-.039	.030	-.172*	-.170*	.209**	.057	.787**	1	.320**	.094
	Sig. (2-tailed)	.186	.623	.703	.028	.030	.007	.470	.000		.000	.232
	N	162	162	162	162	162	162	162	162	162	162	162
Change	Pearson Correlation	-.216**	-.182*	-.097	-.246**	-.242**	-.270**	-.511**	-.219**	.320**	1	.333**
	Sig. (2-tailed)	.006	.020	.220	.002	.002	.001	.000	.005	.000		.000
	N	162	162	162	162	162	162	162	162	162	162	162
Overlap	Pearson Correlation	-.646**	-.386**	-.343**	-.273**	-.917**	.177*	-.711**	-.099	.094	.333**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.024	.000	.208	.232	.000	
	N	162	162	162	162	162	162	162	162	162	162	162

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Appendix E – Data from User Experiment

E5.2 Hand Translation

Correlations

		Not Sure	Difficulty	Confidence	Knowledge	Missed Relevant	False Alarm	M + F	Residual Initial	Final Precision	Change	Overlap
Not Sure	Pearson Correlation	1	.586**	.552**	.355**	.284*	.039	.397**	.117	-.065	-.200	-.242
	Sig. (2-tailed)	.	.000	.000	.008	.037	.779	.003	.399	.642	.148	.078
	N	54	54	54	54	54	54	54	54	54	54	54
Difficulty	Pearson Correlation	.586**	1	.671**	.405**	.178	.152	.447**	.293*	.040	-.325*	-.151
	Sig. (2-tailed)	.000	.	.000	.002	.199	.273	.001	.031	.773	.016	.277
	N	54	54	54	54	54	54	54	54	54	54	54
Confidence	Pearson Correlation	.552**	.671**	1	.478**	.149	.303*	.499**	.429**	.186	-.330*	-.128
	Sig. (2-tailed)	.000	.000	.	.000	.282	.026	.000	.001	.178	.015	.357
	N	54	54	54	54	54	54	54	54	54	54	54
Knowledge	Pearson Correlation	.355**	.405**	.478**	1	.045	.155	.197	.025	-.151	-.273*	-.061
	Sig. (2-tailed)	.008	.002	.000	.	.744	.263	.154	.855	.276	.046	.663
	N	54	54	54	54	54	54	54	54	54	54	54
Missed Relevant	Pearson Correlation	.284*	.178	.149	.045	1	-.236	.556**	.007	-.149	-.121	-.829**
	Sig. (2-tailed)	.037	.199	.282	.744	.	.086	.000	.960	.281	.383	.000
	N	54	54	54	54	54	54	54	54	54	54	54
False Alarm	Pearson Correlation	.039	.152	.303*	.155	-.236	1	.507**	.400**	.189	-.367**	-.066
	Sig. (2-tailed)	.779	.273	.026	.263	.086	.	.000	.003	.172	.006	.636
	N	54	54	54	54	54	54	54	54	54	54	54
M + F	Pearson Correlation	.397**	.447**	.499**	.197	.556**	.507**	1	.389**	.041	-.528**	-.736**
	Sig. (2-tailed)	.003	.001	.000	.154	.000	.000	.	.004	.770	.000	.000
	N	54	54	54	54	54	54	54	54	54	54	54
Residual Initial	Pearson Correlation	.117	.293*	.429**	.025	.007	.400**	.389**	1	.821**	-.200	-.106
	Sig. (2-tailed)	.399	.031	.001	.855	.960	.003	.004	.	.000	.147	.447
	N	54	54	54	54	54	54	54	54	54	54	54
Final Precision	Pearson Correlation	-.065	.040	.186	-.151	-.149	.189	.041	.821**	1	.277*	.095
	Sig. (2-tailed)	.642	.773	.178	.276	.281	.172	.770	.000	.	.043	.494
	N	54	54	54	54	54	54	54	54	54	54	54
Change	Pearson Correlation	-.200	-.325*	-.330*	-.273*	-.121	-.367**	-.528**	-.200	.277*	1	.247
	Sig. (2-tailed)	.148	.016	.015	.046	.383	.006	.000	.147	.043	.	.072
	N	54	54	54	54	54	54	54	54	54	54	54
Overlap	Pearson Correlation	-.242	-.151	-.128	-.061	-.829**	-.066	-.736**	-.106	.095	.247	1
	Sig. (2-tailed)	.078	.277	.357	.663	.000	.636	.000	.447	.494	.072	.
	N	54	54	54	54	54	54	54	54	54	54	54

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Appendix E – Data from User Experiment

E5.3 Machine Translation

Correlations

		Not Sure	Difficulty	Confidence	Knowledge	Missed Relevant	False Alarm	M + F	Residual Initial	Final Precision	Change	Overlap
Not Sure	Pearson Correlation	1	.226	.303*	.204	.411**	-.143	.329*	.105	-.012	-.190	-.362**
	Sig. (2-tailed)	.	.100	.026	.138	.002	.303	.015	.448	.933	.169	.007
	N	54	54	54	54	54	54	54	54	54	54	54
Difficulty	Pearson Correlation	.226	1	.776**	.492**	.080	-.101	-.021	-.008	-.056	-.177	-.078
	Sig. (2-tailed)	.100	.	.000	.000	.566	.469	.880	.954	.687	.201	.574
	N	54	54	54	54	54	54	54	54	54	54	54
Confidence	Pearson Correlation	.303*	.776**	1	.437**	.204	-.102	.113	.032	.019	-.052	-.198
	Sig. (2-tailed)	.026	.000	.	.001	.139	.462	.417	.820	.891	.709	.151
	N	54	54	54	54	54	54	54	54	54	54	54
Knowledge	Pearson Correlation	.204	.492**	.437**	1	.100	-.079	.042	-.098	-.274*	-.407**	-.074
	Sig. (2-tailed)	.138	.000	.001	.	.473	.572	.761	.479	.045	.002	.594
	N	54	54	54	54	54	54	54	54	54	54	54
Missed Relevant	Pearson Correlation	.411**	.080	.204	.100	1	-.251	.679**	.051	-.070	-.248	-.891**
	Sig. (2-tailed)	.002	.566	.139	.473	.	.068	.000	.716	.614	.071	.000
	N	54	54	54	54	54	54	54	54	54	54	54
False Alarm	Pearson Correlation	-.143	-.101	-.102	-.079	-.251	1	.431**	.479**	.330*	-.272*	-.130
	Sig. (2-tailed)	.303	.469	.462	.572	.068	.	.001	.000	.015	.047	.350
	N	54	54	54	54	54	54	54	54	54	54	54
M + F	Pearson Correlation	.329*	-.021	.113	.042	.679**	.431**	1	.472**	.204	-.472**	-.851**
	Sig. (2-tailed)	.015	.880	.417	.761	.000	.001	.	.000	.140	.000	.000
	N	54	54	54	54	54	54	54	54	54	54	54
Residual Initial	Pearson Correlation	.105	-.008	.032	-.098	.051	.479**	.472**	1	.772**	-.246	-.228
	Sig. (2-tailed)	.448	.954	.820	.479	.716	.000	.000	.	.000	.073	.098
	N	54	54	54	54	54	54	54	54	54	54	54
Final Precision	Pearson Correlation	-.012	-.056	.019	-.274*	-.070	.330*	.204	.772**	1	.322*	-.080
	Sig. (2-tailed)	.933	.687	.891	.045	.614	.015	.140	.000	.	.017	.563
	N	54	54	54	54	54	54	54	54	54	54	54
Change	Pearson Correlation	-.190	-.177	-.052	-.407**	-.248	-.272*	-.472**	-.246	.322*	1	.312*
	Sig. (2-tailed)	.169	.201	.709	.002	.071	.047	.000	.073	.017	.	.021
	N	54	54	54	54	54	54	54	54	54	54	54
Overlap	Pearson Correlation	-.362**	-.078	-.198	-.074	-.891**	-.130	-.851**	-.228	-.080	.312*	1
	Sig. (2-tailed)	.007	.574	.151	.594	.000	.350	.000	.098	.563	.021	.
	N	54	54	54	54	54	54	54	54	54	54	54

*. Correlation is significant at the 0.05 level (2-tailed).

** Correlation is significant at the 0.01 level (2-tailed).

Appendix E – Data from User Experiment

E5.4 Original Texts

Correlations

		Not Sure	Difficulty	Confidence	Knowledge	Missed Relevant	False Alarm	M + F	Residual Initial	Final Precision	Change	Overlap
Not Sure	Pearson Correlation	1	.564**	.379**	.464**	.765**	-.562**	.317*	-.025	-.206	-.355**	-.767**
	Sig. (2-tailed)	.	.000	.005	.000	.000	.000	.019	.860	.134	.009	.000
	N	54	54	54	54	54	54	54	54	54	54	54
Difficulty	Pearson Correlation	.564**	1	.378**	.391**	.459**	-.482**	.186	-.025	-.080	-.096	-.429**
	Sig. (2-tailed)	.000	.	.005	.003	.000	.000	.179	.856	.567	.492	.001
	N	54	54	54	54	54	54	54	54	54	54	54
Confidence	Pearson Correlation	.379**	.378**	1	.403**	.317*	-.440**	-.108	-.099	-.066	.080	-.255
	Sig. (2-tailed)	.005	.005	.	.003	.020	.001	.436	.478	.634	.566	.062
	N	54	54	54	54	54	54	54	54	54	54	54
Knowledge	Pearson Correlation	.464**	.391**	.403**	1	.395**	-.478**	.091	.027	-.052	-.019	-.363**
	Sig. (2-tailed)	.000	.003	.003	.	.003	.000	.515	.845	.710	.891	.007
	N	54	54	54	54	54	54	54	54	54	54	54
Missed Relevant	Pearson Correlation	.765**	.459**	.317*	.395**	1	-.583**	.478**	-.085	-.307*	-.384**	-.960**
	Sig. (2-tailed)	.000	.000	.020	.003	.	.000	.000	.539	.024	.004	.000
	N	54	54	54	54	54	54	54	54	54	54	54
False Alarm	Pearson Correlation	-.562**	-.482**	-.440**	-.478**	-.583**	1	.146	.201	.097	-.224	.415**
	Sig. (2-tailed)	.000	.000	.001	.000	.000	.	.293	.145	.484	.103	.002
	N	54	54	54	54	54	54	54	54	54	54	54
M + F	Pearson Correlation	.317*	.186	-.108	.091	.478**	.146	1	.307*	-.048	-.542**	-.557**
	Sig. (2-tailed)	.019	.179	.436	.515	.000	.293	.	.024	.733	.000	.000
	N	54	54	54	54	54	54	54	54	54	54	54
Residual Initial	Pearson Correlation	-.025	-.025	-.099	.027	-.085	.201	.307*	1	.770**	-.211	-.004
	Sig. (2-tailed)	.860	.856	.478	.845	.539	.145	.024	.	.000	.125	.979
	N	54	54	54	54	54	54	54	54	54	54	54
Final Precision	Pearson Correlation	-.206	-.080	-.066	-.052	-.307*	.097	-.048	.770**	1	.369**	.264
	Sig. (2-tailed)	.134	.567	.634	.710	.024	.484	.733	.000	.	.006	.054
	N	54	54	54	54	54	54	54	54	54	54	54
Change	Pearson Correlation	-.355**	-.096	.080	-.019	-.384**	-.224	-.542**	-.211	.369**	1	.487**
	Sig. (2-tailed)	.009	.492	.566	.891	.004	.103	.000	.125	.006	.	.000
	N	54	54	54	54	54	54	54	54	54	54	54
Overlap	Pearson Correlation	-.767**	-.429**	-.255	-.363**	-.960**	.415**	-.557**	-.004	.264	.487**	1
	Sig. (2-tailed)	.000	.001	.062	.007	.000	.002	.000	.979	.054	.000	.
	N	54	54	54	54	54	54	54	54	54	54	54

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Appendix E – Data from User Experiment

E5.5 Topic 1

Correlations

		Not Sure	Difficulty	Confidence	Knowledge	Missed Relevant	False Alarm	M + F	Change	Overlap
Not Sure	Pearson Correlation	1	.666**	.377	.615**	.812**	-.622**	.623**	-.886**	-.843**
	Sig. (2-tailed)	.	.000	.053	.001	.000	.001	.001	.000	.000
	N	27	27	27	27	27	27	27	27	27
Difficulty	Pearson Correlation	.666**	1	.522**	.808**	.825**	-.735**	.507**	-.665**	-.760**
	Sig. (2-tailed)	.000	.	.005	.000	.000	.000	.007	.000	.000
	N	27	27	27	27	27	27	27	27	27
Confidence	Pearson Correlation	.377	.522**	1	.706**	.498**	-.427*	.326	-.437*	-.461*
	Sig. (2-tailed)	.053	.005	.	.000	.008	.026	.097	.023	.015
	N	27	27	27	27	27	27	27	27	27
Knowledge	Pearson Correlation	.615**	.808**	.706**	1	.713**	-.630**	.444*	-.635**	-.648**
	Sig. (2-tailed)	.001	.000	.000	.	.000	.000	.020	.000	.000
	N	27	27	27	27	27	27	27	27	27
Missed Relevant	Pearson Correlation	.812**	.825**	.498**	.713**	1	-.814**	.708**	-.826**	-.963**
	Sig. (2-tailed)	.000	.000	.008	.000	.	.000	.000	.000	.000
	N	27	27	27	27	27	27	27	27	27
False Alarm	Pearson Correlation	-.622**	-.735**	-.427*	-.630**	-.814**	1	-.166	.671**	.650**
	Sig. (2-tailed)	.001	.000	.026	.000	.000	.	.408	.000	.000
	N	27	27	27	27	27	27	27	27	27
M + F	Pearson Correlation	.623**	.507**	.326	.444*	.708**	-.166	1	-.586**	-.846**
	Sig. (2-tailed)	.001	.007	.097	.020	.000	.408	.	.001	.000
	N	27	27	27	27	27	27	27	27	27
Change	Pearson Correlation	-.886**	-.665**	-.437*	-.635**	-.826**	.671**	-.586**	1	.859**
	Sig. (2-tailed)	.000	.000	.023	.000	.000	.000	.001	.	.000
	N	27	27	27	27	27	27	27	27	27
Overlap	Pearson Correlation	-.843**	-.760**	-.461*	-.648**	-.963**	.650**	-.846**	.859**	1
	Sig. (2-tailed)	.000	.000	.015	.000	.000	.000	.000	.000	.
	N	27	27	27	27	27	27	27	27	27

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Appendix E – Data from User Experiment

E5.6 Topic 2

Correlations

		Not Sure	Difficulty	Confidence	Knowledge	Missed Relevant	False Alarm	M + F	Change	Overlap
Not Sure	Pearson Correlation	1	.717**	.413*	.455*	.821**	-.407*	.787**	.516**	-.819**
	Sig. (2-tailed)	.	.000	.032	.017	.000	.035	.000	.006	.000
	N	27	27	27	27	27	27	27	27	27
Difficulty	Pearson Correlation	.717**	1	.477*	.245	.655**	-.370	.612**	.246	-.639**
	Sig. (2-tailed)	.000	.	.012	.219	.000	.057	.001	.217	.000
	N	27	27	27	27	27	27	27	27	27
Confidence	Pearson Correlation	.413*	.477*	1	.298	.427*	-.260	.392*	.186	-.367
	Sig. (2-tailed)	.032	.012	.	.132	.026	.190	.043	.353	.060
	N	27	27	27	27	27	27	27	27	27
Knowledge	Pearson Correlation	.455*	.245	.298	1	.299	-.133	.292	.259	-.298
	Sig. (2-tailed)	.017	.219	.132	.	.130	.508	.140	.191	.131
	N	27	27	27	27	27	27	27	27	27
Missed Relevant	Pearson Correlation	.821**	.655**	.427*	.299	1	-.511**	.953**	.547**	-.985**
	Sig. (2-tailed)	.000	.000	.026	.130	.	.006	.000	.003	.000
	N	27	27	27	27	27	27	27	27	27
False Alarm	Pearson Correlation	-.407*	-.370	-.260	-.133	-.511**	1	-.226	-.607**	.394*
	Sig. (2-tailed)	.035	.057	.190	.508	.006	.	.257	.001	.042
	N	27	27	27	27	27	27	27	27	27
M + F	Pearson Correlation	.787**	.612**	.392*	.292	.953**	-.226	1	.406*	-.977**
	Sig. (2-tailed)	.000	.001	.043	.140	.000	.257	.	.036	.000
	N	27	27	27	27	27	27	27	27	27
Change	Pearson Correlation	.516**	.246	.186	.259	.547**	-.607**	.406*	1	-.519**
	Sig. (2-tailed)	.006	.217	.353	.191	.003	.001	.036	.	.005
	N	27	27	27	27	27	27	27	27	27
Overlap	Pearson Correlation	-.819**	-.639**	-.367	-.298	-.985**	.394*	-.977**	-.519**	1
	Sig. (2-tailed)	.000	.000	.060	.131	.000	.042	.000	.005	.
	N	27	27	27	27	27	27	27	27	27

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Appendix E – Data from User Experiment

E5.7 Topic 3

Correlations

		Not Sure	Difficulty	Confidence	Knowledge	Missed Relevant	False Alarm	M + F	Change	Overlap
Not Sure	Pearson Correlation	1	.673**	.731**	.447*	.805**	-.349	.352	-.102	-.769**
	Sig. (2-tailed)	.	.000	.000	.020	.000	.075	.072	.613	.000
	N	27	27	27	27	27	27	27	27	27
Difficulty	Pearson Correlation	.673**	1	.712**	.580**	.646**	-.215	.339	-.016	-.592**
	Sig. (2-tailed)	.000	.	.000	.002	.000	.280	.084	.936	.001
	N	27	27	27	27	27	27	27	27	27
Confidence	Pearson Correlation	.731**	.712**	1	.398*	.606**	-.373	.169	.308	-.529**
	Sig. (2-tailed)	.000	.000	.	.040	.001	.056	.400	.119	.005
	N	27	27	27	27	27	27	27	27	27
Knowledge	Pearson Correlation	.447*	.580**	.398*	1	.372	-.099	.217	-.149	-.372
	Sig. (2-tailed)	.020	.002	.040	.	.056	.623	.276	.458	.056
	N	27	27	27	27	27	27	27	27	27
Missed Relevant	Pearson Correlation	.805**	.646**	.606**	.372	1	-.307	.549**	-.050	-.922**
	Sig. (2-tailed)	.000	.000	.001	.056	.	.120	.003	.803	.000
	N	27	27	27	27	27	27	27	27	27
False Alarm	Pearson Correlation	-.349	-.215	-.373	-.099	-.307	1	.627**	-.484*	-.029
	Sig. (2-tailed)	.075	.280	.056	.623	.120	.	.000	.011	.886
	N	27	27	27	27	27	27	27	27	27
M + F	Pearson Correlation	.352	.339	.169	.217	.549**	.627**	1	-.466*	-.780**
	Sig. (2-tailed)	.072	.084	.400	.276	.003	.000	.	.014	.000
	N	27	27	27	27	27	27	27	27	27
Change	Pearson Correlation	-.102	-.016	.308	-.149	-.050	-.484*	-.466*	1	.237
	Sig. (2-tailed)	.613	.936	.119	.458	.803	.011	.014	.	.233
	N	27	27	27	27	27	27	27	27	27
Overlap	Pearson Correlation	-.769**	-.592**	-.529**	-.372	-.922**	-.029	-.780**	.237	1
	Sig. (2-tailed)	.000	.001	.005	.056	.000	.886	.000	.233	.
	N	27	27	27	27	27	27	27	27	27

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Appendix E – Data from User Experiment

E5.8 Topic 4

Correlations

		Not Sure	Difficulty	Confidence	Knowledge	Missed Relevant	False Alarm	M + F	Change	Overlap
Not Sure	Pearson Correlation	1	.636**	.681**	.504**	.553**	-.233	.086	-.596**	-.557**
	Sig. (2-tailed)	.	.000	.000	.007	.003	.242	.669	.001	.003
	N	27	27	27	27	27	27	27	27	27
Difficulty	Pearson Correlation	.636**	1	.965**	.591**	.283	-.071	.085	-.389*	-.238
	Sig. (2-tailed)	.000	.	.000	.001	.152	.726	.675	.045	.232
	N	27	27	27	27	27	27	27	27	27
Confidence	Pearson Correlation	.681**	.965**	1	.620**	.331	-.051	.125	-.409*	-.276
	Sig. (2-tailed)	.000	.000	.	.001	.092	.800	.535	.034	.163
	N	27	27	27	27	27	27	27	27	27
Knowledge	Pearson Correlation	.504**	.591**	.620**	1	.400*	.000	.202	-.471*	-.459*
	Sig. (2-tailed)	.007	.001	.001	.	.039	1.000	.312	.013	.016
	N	27	27	27	27	27	27	27	27	27
Missed Relevant	Pearson Correlation	.553**	.283	.331	.400*	1	.068	.562**	-.591**	-.873**
	Sig. (2-tailed)	.003	.152	.092	.039	.	.735	.002	.001	.000
	N	27	27	27	27	27	27	27	27	27
False Alarm	Pearson Correlation	-.233	-.071	-.051	.000	.068	1	.863**	-.260	-.348
	Sig. (2-tailed)	.242	.726	.800	1.000	.735	.	.000	.191	.075
	N	27	27	27	27	27	27	27	27	27
M + F	Pearson Correlation	.086	.085	.125	.202	.562**	.863**	1	-.514**	-.730**
	Sig. (2-tailed)	.669	.675	.535	.312	.002	.000	.	.006	.000
	N	27	27	27	27	27	27	27	27	27
Change	Pearson Correlation	-.596**	-.389*	-.409*	-.471*	-.591**	-.260	-.514**	1	.703**
	Sig. (2-tailed)	.001	.045	.034	.013	.001	.191	.006	.	.000
	N	27	27	27	27	27	27	27	27	27
Overlap	Pearson Correlation	-.557**	-.238	-.276	-.459*	-.873**	-.348	-.730**	.703**	1
	Sig. (2-tailed)	.003	.232	.163	.016	.000	.075	.000	.000	.
	N	27	27	27	27	27	27	27	27	27

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Appendix E – Data from User Experiment

E5.9 Topic 5

Correlations

		Not Sure	Difficulty	Confidence	Knowledge	Missed Relevant	False Alarm	M + F	Change	Overlap
Not Sure	Pearson Correlation	1	.452*	.549**	.327	.484*	-.265	.406*	-.433*	-.482*
	Sig. (2-tailed)	.	.018	.003	.096	.011	.181	.035	.024	.011
	N	27	27	27	27	27	27	27	27	27
Difficulty	Pearson Correlation	.452*	1	.762**	.425*	.045	-.063	.020	-.155	-.061
	Sig. (2-tailed)	.018	.	.000	.027	.823	.756	.922	.439	.763
	N	27	27	27	27	27	27	27	27	27
Confidence	Pearson Correlation	.549**	.762**	1	.501**	.112	-.034	.107	-.228	-.126
	Sig. (2-tailed)	.003	.000	.	.008	.578	.867	.594	.253	.531
	N	27	27	27	27	27	27	27	27	27
Knowledge	Pearson Correlation	.327	.425*	.501**	1	-.130	.006	-.140	-.267	.131
	Sig. (2-tailed)	.096	.027	.008	.	.519	.977	.485	.178	.515
	N	27	27	27	27	27	27	27	27	27
Missed Relevant	Pearson Correlation	.484*	.045	.112	-.130	1	-.421*	.901**	-.564**	-.982**
	Sig. (2-tailed)	.011	.823	.578	.519	.	.029	.000	.002	.000
	N	27	27	27	27	27	27	27	27	27
False Alarm	Pearson Correlation	-.265	-.063	-.034	.006	-.421*	1	.015	.022	.268
	Sig. (2-tailed)	.181	.756	.867	.977	.029	.	.940	.915	.177
	N	27	27	27	27	27	27	27	27	27
M + F	Pearson Correlation	.406*	.020	.107	-.140	.901**	.015	1	-.612**	-.955**
	Sig. (2-tailed)	.035	.922	.594	.485	.000	.940	.	.001	.000
	N	27	27	27	27	27	27	27	27	27
Change	Pearson Correlation	-.433*	-.155	-.228	-.267	-.564**	.022	-.612**	1	.613**
	Sig. (2-tailed)	.024	.439	.253	.178	.002	.915	.001	.	.001
	N	27	27	27	27	27	27	27	27	27
Overlap	Pearson Correlation	-.482*	-.061	-.126	.131	-.982**	.268	-.955**	.613**	1
	Sig. (2-tailed)	.011	.763	.531	.515	.000	.177	.000	.001	.
	N	27	27	27	27	27	27	27	27	27

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Appendix E – Data from User Experiment

E5.10 Topic 6

Correlations

		Not Sure	Difficulty	Confidence	Knowledge	Missed Relevant	False Alarm	M + F	Change	Overlap
Not Sure	Pearson Correlation	1	.508**	.477*	.192	.562**	-.648**	-.034	.032	-.560**
	Sig. (2-tailed)	.	.007	.012	.337	.002	.000	.865	.875	.002
	N	27	27	27	27	27	27	27	27	27
Difficulty	Pearson Correlation	.508**	1	.640**	.467*	.096	-.306	-.238	-.032	-.116
	Sig. (2-tailed)	.007	.	.000	.014	.634	.121	.232	.875	.564
	N	27	27	27	27	27	27	27	27	27
Confidence	Pearson Correlation	.477*	.640**	1	.539**	.434*	-.434*	.052	-.117	-.436*
	Sig. (2-tailed)	.012	.000	.	.004	.024	.024	.798	.562	.023
	N	27	27	27	27	27	27	27	27	27
Knowledge	Pearson Correlation	.192	.467*	.539**	1	-.022	-.201	-.269	.005	-.003
	Sig. (2-tailed)	.337	.014	.004	.	.912	.314	.175	.981	.989
	N	27	27	27	27	27	27	27	27	27
Missed Relevant	Pearson Correlation	.562**	.096	.434*	-.022	1	-.684**	.496**	-.213	-.958**
	Sig. (2-tailed)	.002	.634	.024	.912	.	.000	.008	.287	.000
	N	27	27	27	27	27	27	27	27	27
False Alarm	Pearson Correlation	-.648**	-.306	-.434*	-.201	-.684**	1	.294	.102	.525**
	Sig. (2-tailed)	.000	.121	.024	.314	.000	.	.136	.612	.005
	N	27	27	27	27	27	27	27	27	27
M + F	Pearson Correlation	-.034	-.238	.052	-.269	.496**	.294	1	-.157	-.631**
	Sig. (2-tailed)	.865	.232	.798	.175	.008	.136	.	.435	.000
	N	27	27	27	27	27	27	27	27	27
Change	Pearson Correlation	.032	-.032	-.117	.005	-.213	.102	-.157	1	.248
	Sig. (2-tailed)	.875	.875	.562	.981	.287	.612	.435	.	.213
	N	27	27	27	27	27	27	27	27	27
Overlap	Pearson Correlation	-.560**	-.116	-.436*	-.003	-.958**	.525**	-.631**	.248	1
	Sig. (2-tailed)	.002	.564	.023	.989	.000	.005	.000	.213	.
	N	27	27	27	27	27	27	27	27	27

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).